

OM of: US-08-711-417C-165 to: Pending\_Patents\_AA\_Main:\* out\_format : pfs

Date: Aug 28, 2002 10:13 AM

About: Results were produced by the GenCore software, version 4.5,  
Copyright (c) 1993-2000 Compugen Ltd.

# Command line parameters:

```
-MODEL=frame+np2.model -DEV=xlp
-Q/cgn2_1/USPTO.spool/6228611/runat_28082002_100211_13594/app_query.fasta_1.1639
-DB=Pending_Patents_AA_Main -OFM=fastan -SUFFIX=rapm
-GAPOP=12.000 -GAPEXT=4.000 -MINMATCH=0.100 -LOOPEX=0.000
-LOOPEXT=0.000 -GAPOP=4.500 -GAPEXT=0.050 -GAPOP=10.000
-YGAPEXT=0.500 -DELOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000
-MATRIX=blosum62 -TRANS=human40.cdi -LIST=45 -DOCLIGN=200
-THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pfs -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=6228611.ecn1_1.784 -NCPU=6 -ICPU=3 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -NO_XLUPXY -WAIT -THREADS=1
```

## Search information block:

```
Query: US-08-711-417C-165
Query length: 1551
Database: Pending_Patents_AA_Main:*
Database sequence: 351980561
Database length: 351980561
Search time (sec): 370.380000
```

## score\_list:

Sequence	Strd Orig	Zscore	Escore	Len	Documentation
/cgn2_6/ptodata/2/paa/US090_COMB.ppt:US99-02559-25	2467.00	2936.88	1.3e-155	461	
/cgn2_6/ptodata/2/paa/US090_COMB.ppt:US-09-019-348-25	2467.00	2936.88	1.3e-155	461	
/cgn2_6/ptodata/2/paa/US090_COMB.ppt:US-09-019-348A-25	2467.00	2936.88	1.3e-155	461	
/cgn2_6/ptodata/2/paa/US097_COMB.ppt:US-09-755-830-38	2467.00	2936.88	1.3e-155	461	
/cgn2_6/ptodata/2/paa/US090_COMB.ppt:US-09-019-348-27	2437.00	2899.93	1.4e-153	518	
/cgn2_6/ptodata/2/paa/US090_COMB.ppt:US-09-019-348-27	2437.00	2899.93	1.4e-153	518	
/cgn2_6/ptodata/2/paa/US090_COMB.ppt:US-09-019-348A-27	2437.00	2899.93	1.4e-153	518	
/cgn2_6/ptodata/2/paa/US097_COMB.ppt:US-09-755-830-40	2437.00	2899.93	1.4e-153	518	
/cgn2_6/ptodata/2/paa/US090_COMB.ppt:US99-04224-7	2426.50	2887.42	6.7e-153	517	
/cgn2_6/ptodata/2/paa/US092_COMB.ppt:US-09-259-389-7	2426.50	2887.42	6.7e-153	517	
/cgn2_6/ptodata/2/paa/US081_COMB.ppt:US-08-121-434-5	2422.00	2881.14	1.4e-152	568	
/cgn2_6/ptodata/2/paa/US090_COMB.ppt:US99-02559-22	2207.50	2626.84	2.4e-138	470	
/cgn2_6/ptodata/2/paa/US082_COMB.ppt:US-08-238-212A-153	2207.50	2626.84	2.4e-138	470	
/cgn2_6/ptodata/2/paa/US084_COMB.ppt:US-08-465-5908-153	2207.50	2626.84	2.4e-138	470	
/cgn2_6/ptodata/2/paa/US087_COMB.ppt:US-08-733-622A-22	2207.50	2626.84	2.4e-138	470	
/cgn2_6/ptodata/2/paa/US090_COMB.ppt:US-09-019-348-22	2207.50	2626.84	2.4e-138	470	
/cgn2_6/ptodata/2/paa/US090_COMB.ppt:US-09-019-348A-22	2207.50	2626.84	2.4e-138	470	
/cgn2_6/ptodata/2/paa/US097_COMB.ppt:US-09-755-830-13	2207.50	2626.84	2.4e-138	470	
/cgn2_6/ptodata/2/paa/US097_COMB.ppt:US99-02559-26	1963.00	2335.71	4.3e-122	432	
/cgn2_6/ptodata/2/paa/US090_COMB.ppt:US-09-019-348-26	1963.00	2335.71	4.3e-122	432	
/cgn2_6/ptodata/2/paa/US097_COMB.ppt:US-09-755-830-39	1963.00	2335.71	4.3e-122	432	
/cgn2_6/ptodata/2/paa/US090_COMB.ppt:US99-02559-24	1923.50	2288.57	1.8e-119	431	
/cgn2_6/ptodata/2/paa/US090_COMB.ppt:US-09-019-348-24	1923.50	2288.57	1.8e-119	431	
/cgn2_6/ptodata/2/paa/US097_COMB.ppt:US-09-019-348A-24	1923.50	2288.57	1.8e-119	431	
/cgn2_6/ptodata/2/paa/US097_COMB.ppt:US-09-755-830-37	1920.50	2284.99	2.9e-119	431	
/cgn2_6/ptodata/2/paa/US090_COMB.ppt:US99-02559-23	1740.00	2071.92	2.8e-107	334	
/cgn2_6/ptodata/2/paa/US090_COMB.ppt:US-09-019-348-23	1740.00	2071.92	2.8e-107	334	
/cgn2_6/ptodata/2/paa/US097_COMB.ppt:US-09-755-830-43	1740.00	2071.92	2.8e-107	334	
/cgn2_6/ptodata/2/paa/US097_COMB.ppt:US99-02559-28	1681.00	1999.98	2.4e-103	390	
/cgn2_6/ptodata/2/paa/US090_COMB.ppt:US-09-019-348-28	1681.00	1999.98	2.4e-103	390	
/cgn2_6/ptodata/2/paa/US097_COMB.ppt:US-09-755-830-41	1681.00	1999.98	2.4e-103	390	
/cgn2_6/ptodata/2/paa/US090_COMB.ppt:US99-02559-29	1625.00	1933.47	1.3e-99	376	
/cgn2_6/ptodata/2/paa/US090_COMB.ppt:US-09-019-348-29	1625.00	1933.47	1.3e-99	376	
/cgn2_6/ptodata/2/paa/US097_COMB.ppt:US-09-019-348A-29	1625.00	1933.47	1.3e-99	376	
/cgn2_6/ptodata/2/paa/US087_COMB.ppt:US-08-733-622B-27	1556.00	1852.94	4.7e-95	310	

```
/cgn2_6/ptodata/2/paa/US090_COMB.ppt:US99-04224-2 + 1371.50 1627.55 9.9e-83 52
/cgn2_6/ptodata/2/paa/US092_COMB.ppt:US-09-259-389-2 + 1371.50 1627.55 9.9e-83 52
/cgn2_6/ptodata/2/paa/US090_COMB.ppt:US99-04224-6 + 1369.00 1624.56 1.5e-82 52
/cgn2_6/ptodata/2/paa/US092_COMB.ppt:US-09-259-389-6 + 1369.00 1624.56 1.5e-82 52
/cgn2_6/ptodata/2/paa/US060_COMB.ppt:US-60-243-468-1020 + 1309.00 1552.25 1.4e-78 5
```

seq\_name: /cgn2\_6/ptodata/2/paa/PCTUS\_COMB.ppt:US99-02559-25

```
seq_documentation_block:
; Sequence 25, Application PC/TUS9902559
; GENERAL INFORMATION:
; APPLICANT: Shiseido Co., Ltd.
; TITLE OF INVENTION: AIOLOS GENE
; FILE REFERENCE: 10287/031W01
; CURRENT APPLICATION NUMBER: PCT/US99/02559
; CURRENT FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: US 09/019,348
; EARLIER FILING DATE: 1998-02-05
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 25
; LENGTH: 461
; TYPE: PRT
; ORGANISM: Mus musculus
PCT-US99-02559-25
```

alignment\_scores:  
Quality: 2467.00 Length: 461  
Ratio: 5.351 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

## alignment\_block:

```
US-08-711-417C-165 x PCT-US99-02559-25
Align seg 1/1 to: PCT-US99-02559-25 from: 1 to: 461

166 AATGTTAAAGTAGAGCTCAGAGTGTAGAGAGAAATGGCGTCTGTGA 215
|||||
1 AsnValLysValGluThrGlnSerAspGluAsnGlyArgAlaCysG1 17

216 AATGAATGGGGAAGAATGTCGGAGGATTACGAATGCTGATCCCTCGG 265
|||||
17 uMetAsnGlyGluGlyCysAlaGluAspLeuArgMetLeuAspAlaSerG 34

266 GAGAGAAATGAATGGCTCCACAGGACCAAGCAGCTCGGCTTTCGCG 315
|||||
34 lYGLuLysMetAsnGlySerHisArgAspGlnGlySerSerAlaLeuSer 50

316 GGAGTTGGAGGACATCGACTTCCTAACGGAAACCTAAGTGTGATATCG 365
|||||
51 GlyValGlyGlyIleArgLeuProAsnGlyLysLeuLysCysAspIleCy 67

366 TGGGATCATTTTCATCGGGCCCAATGCTCATGTTTCAAAAAGAGGCC 415
|||||
67 sGlyIleLeuLysGlyGlyProAsnValLeuMetValHisLysArgSerH 84

416 ACAGTGGAGAACGCCCTTCAGTGCATCAGTCCGGGGCTCATTCACC 465
|||||
84 lThrGlyGluArgProPheGlnCysAsnGlnCysGlyAlaSerPheThr 100

466 CAGAGGGCAACCTGCTCCGGCCACATCAAGCTGCATTCGGGGGAGAAGCC 515
|||||
101 GlnLysGlyAsnLeuLeuArgHisIleLysLeuHisSerGlyGluLysPr 117

516 CTTCAATGCGCACCTCTGCAACTACGCTGCGCGGAGGAGGCCCTCA 565
|||||
117 oPhelyCysHisLeuLysCysAsnTyrAlaCysArgArgAspAlaLeuT 134

566 CTGGCCACCTGAGGAGCAGCTCCGTTGGTAAACCTCAAAATGTGGATAT 615
|||||
134 hrGlyHisLeuArgThrHisSerValGlyLysProHisLysCysGlyTyr 150
```

```
616 TGTGGCCGAGCTATTAACAGCGAAGCTCTTTAGAGGAACATAAGAGCG 665
|||||
151 CysGlyArgSerTyrLysGlnArgThrSerLeuGluHisLysGluAr 167
666 CTGCCCAACTACTTGGAAAGCATGGCCCTTCGGGCACACTGTACCCAG 715
|||||
167 gCysHisAsnTyrLeuGluSerMetGlyLeuProGlyThrLeuTyrProV 184
716 TCATTAAAGAAGAACTAAGCAGCAGTGAATGGCAGAGACCTGTGCAAG 765
|||||
184 alileLysGluGluThrLysHisSerGluMetAlaGluAspLeuLys 200
766 ATAGGATCAGAGAGATCTCTCGCTGGACAGACTAGCAAGTAAGTGC 815
|||||
201 IleGlySerGluArgSerLeuValLeuAspArgLeuAlaSerAsnValAl 217
816 CAACGTAGAGCTCTATGCCCTCAGAAATTTCTTGGGACAGGCGCTGT 865
|||||
217 aLysArgLysSerSerMetProGlnLysPheLeuGlyAspLysGlyLeu 234
866 CGCACAGCGCCTACGACAGTGCACGTACGAGAGGAGAACCAATGATG 915
|||||
234 erAspThrProTyrAspSerAlaThrTyrGluLysGluAsnGluMetMet 250
916 AGTCCCGCTGATGACCAAGCCATCAACAGCCATCACTACTGCGG 965
|||||
251 LysSerHisValMetAspGlnAlaIleAsnAsnAlaIleAsnTyrLeuGl 267
966 GCGCGAGTCCCTGCGCCGCTGGTCGACAGCGCCCGCGCGGTCCGAGG 1015
|||||
267 yAlaGluSerLeuArgProLeuValGlnThrProProGlyGlySerGluV 284
1016 TGTGCCCGCTCATACGCCGATGTACCACTGCACAGCGCTCGGAGGGC 1065
|||||
284 alValProValIleSerProMetTyrGlnLeuHisArgArgSerGluGly 300
1066 ACCCGCGCTCCAACCACTCGGCCAGGACAGCGCGGTGGAGTACTGCT 1115
|||||
301 ThrProArgSerAsnHisSerAlaGlnAspSerAlaValGluTyrLeuLe 317
1116 GTGCTCTCCAAAGGCCAAGTTGGTCCCTCGCGAGCGCGAGGCTCCCGA 1165
|||||
317 uLeuLeuSerLysAlaLysLeuValProSerGluArgGluAlaSerPro 334
1166 GCAACAGCTGCCAAGACTCCAGCACCGAGAGCAACCAACAGGAGGAG 1215
|||||
334 erAsnSerCysGlnAspSerThrAspThrGluSerAsnAsnGluGluGln 350
1216 CGGAGCGGTCTTATCTACCTGACCAACACATCGCCCGAGCGCGCAAG 1265
|||||
351 ArgSerGlyLeuIleTyrLeuThrAsnHisIleAlaArgAlaGlnAr 367
1266 CGTGTGCTCAAGGAGGACCGCGCTACGACCTGCTGCGCGCGCT 1315
|||||
367 gValSerLeuLysGluGluHisArgAlaTyrAspLeuLeuArgAlaAla 384
1316 CCAGAACTCGCAGGACCGCTCCCGTGGTCAGCACCAGCGGGGAGCAG 1365
|||||
384 erGluAsnSerGlnAspAlaLeuArgValValSerThrSerGlyGluGln 400
1366 ATGAGGTGTACAAGTGGGAACACTCGCGGTGCTCTTCCTGGATCAGT 1415
|||||
401 MetLysValTyrLysCysGluHisCysArgValLeuPheLeuAspHisVa 417
1416 CATGTACACCATCCACATGGGCTGCCAGCGCTTCGTGATCCTTTGAGT 1465
|||||
417 lMetTyrThrIleHisMetGlyCysHisGlyPheArgAspPropheGluC 434
1466 GCAACATGTGCGGCTACCAAGCCAGGACCGGTAGAGTCTCTCGTCGAC 1515
|||||
434 ysAsnMetCysGlyTyrHisSerGlnAspArgTyrGluPheSerSerHis 450
1516 ATAACGCGAGGGGACCGCTTCCACATGAGC 1548
```

```
|||||
451 IleThrArgGlyGluHisArgPheHisMetSer 461
seq_name: /cqn2_6/ptodata/2/paa/US090_COMB.pep:US-09-019-348-25

seq_documentation_block:
; Sequence 25, Application US/09019348
; GENERAL INFORMATION:
; APPLICANT: Georgopoulos, Katia
; APPLICANT: Morgan, Bruce
; TITLE OF INVENTION: AIOLOS GENE
; FILE REFERENCE: 10287/031001
; CURRENT APPLICATION NUMBER: US/09/019,348
; CURRENT FILING DATE: 1998-02-05
; EARLIER APPLICATION NUMBER: US 08/733,622
; EARLIER FILING DATE: 1996-10-17
; EARLIER APPLICATION NUMBER: US 60/017,646
; EARLIER FILING DATE: 1996-05-14
; EARLIER APPLICATION NUMBER: US 60/005,527
; EARLIER FILING DATE: 1995-10-18
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 25
; LENGTH: 461
; TYPE: PRT
; ORGANISM: Mus musculus
; US-09-019-348-25

alignment_scores:
Quality: 2467.00 Length: 461
Ratio: 5.351 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-711-417C-165 x US-09-019-348-25 ..

Align seg 1/1 to: US-09-019-348-25 from: 1 to: 461

166 AATGTTAAAGTAGAGACTCAGAGTGTAGAGAAATGGCGTGCCTGTGA 215
1 AsnValLysValGluThrGlnSerAspGluGluAsnGlyArgAlaCysGl 17
216 AATGAATGGGGAAGAATGTCCGGAGGATTTACGAATGCTTGTATGCCCTCGG 265
17 uMetAsnGlyGluGluCysAlaGluAspLeuArgMetLeuAspAlaSerG 34
266 GAGAGAAATGAATGGCTCCACAGGACCAAGGACGCTCGGCTTTGTGCG 315
34 lyGluLysMetAsnGlySerHisArgAspGlnGlySerSerAlaLeuSer 50
316 GGAGTTGGAGGCATTCGACTTCCTTAACGGAAACTAAAGTGTGATATCTG 365
51 GlyValGlyGlyIleArgLeuProAsnGlyLysLeuLysCysAspIleCy 67
366 TGGGATCATTTGCATCGGCCCAATGCTCATGTTTCACAAAAGAGCC 415
67 sGlyIleCysIleGlyProAsnValLeuMetValHisLysArgSerH 84
416 ACAGTGAGAACGGCCCTTCCAGTGCAACTCAGTGCAGGCGGCGCTCATTCAC 465
84 isThrGlyGluArgProPheGlnCysAsnGlnCysGlyAlaSerPheThr 100
466 CAGAAAGGCAACCTGCTCCGGCAGCATCAAGCTGCTTCGCGGGGAGAGCC 515
101 GlnLysGlyAsnLeuLeuArgHisIleLysLeuHisSerGlyGluLysPr 117
516 CTTCAAATGCCACTCTGCAACTACGCTCCCGCGGAGGAGCGCCCTCA 565
117 oPheLysCysHisLeuCysAsnTyrAlaCysArgArgAspAlaLeuT 134
566 CTGCCACCTGAGGACGCACTCCGTTGGTAAACCTCACAAATGTGGATAT 615
|||||
```

```

134 hrGlyHisLeuArgThrHisSerValGlyLysProHisLysCysGlyTyr 150
616 TGTGGCCGAGCTATAAAGCAGCAACGCTCTTTAGAGGAACATAAGAGCG 665
151 CysGlyArgSerTyrLysGlnArgThrSerLeuGluGluHisLysGluAr 167
666 CTGCCCAACTACTTGGAAAGCATGGGCTTCCGGGACACACTGTACCCAG 715
167 gCysHisAsnTyrLeuGluSerMetGlyLeuProGlyThrLeuTyrProV 184
716 TCATTAAAGAGAACTAAGCACAGTGAATGGCAGAGACCTGTGCAAG 765
184 alileLysGluGluThrLysHisSerGluMetAlaGluAspLeuCysLys 200
766 ATAGATCAGAGAGATCTCTCGTGTGACAGACTAGCAAGTAAATGTCG 815
201 IleGlySerGluArgSerLeuValLeuAspArgLeuAlaSerAsnValAl 217
816 CAACGCTAAGAGCTCTATGCTCCAGAAATTTCTTGGGACAAGGCGCTGT 865
217 aLysArgLysSerSerMetProGlnLysPheLeuGlyAspLysGlyLeuS 234
866 CGGACACGCCCTCAGACAGTGCCTCCAGTACGAGAGGAGAGAAATGATG 915
234 erAspThrProTyrAspSerAlaThrTyrGluLysGluAsnGluMetMet 950
916 AAGTCCACAGGTGATGGACCAAGCCATCAACAGCCATCAACTACCTGG 965
251 LysSerHisValMetAspGlnAlaIleAsnAsnAlaIleAsnTyrLeuGl 267
966 GGCCGAGTCCCTCGGCCCTGTGTGACAGACGCCCGCGGCTTCGAGG 1015
267 yAlaGluSerLeuArgProLeuValGlnThrProGlyGlySerGluV 284
1016 TGTTCGCCGTCATCAGCCGATGTACAGCTGCACAGGCGCTCGAGGCG 1065
284 alValProValIleSerProMetTyrGlnLeuHisArgArgSerGluGly 300
1066 ACCCGCGCTCCCAACCACTCGGCCAGGACAGCGCTGGAGCGCTCCCGA 1115
301 ThrProArgSerAsnHisSerAlaGlnAspSerAlaValGluTyrLeuLe 317
1116 GCTGCTCTCCAGGCCAAGTTGGTGGCTTCGAGCGCGAGCGCTCCCGA 1165
317 uLeuLeuSerLysAlaLysLeuValProSerGluArgGluAlaSerProS 334
1166 GCAACAGCTGCCAAGACTCCAGGACACCGAGAGCAACACAGGAGGAGC 1215
334 erAsnSerCysGlnAspSerThrAspThrGluSerAsnAsnGluGluGln 350
1216 CGCAGCGGTCTTATCTACTGACCAACCACTCGCGGTCGACGCGCGCAACG 1265
351 ArgSerGlyLeuIleTyrLeuThrAsnHisIleAlaArgArgAlaGlnAr 367
1266 CBTGTGCTTCAAGGAGGAGCAGCGGCTTACAGCTGCTCGCGCGCGCT 1315
367 gValSerLeuLysGluGluHisArgAlaTyrAspLeuArgAlaAlas 384
1316 CCGAGACTCCGAGACCGGCTCGCGGTGTCAGCACACGCGGAGGAGCAG 1365
384 erGluAsnSerGlnAspAlaLeuArgValValSerThrSerGlyGluGln 400
1366 ATGAAGTGTACAAGTGGGAACACTCGCGGTGCTCTCTCTGATACAGT 1415
401 MetLysValTyrLysCysGluHisCysArgValLeuPheLeuAspHisVa 417
1416 CATGTACACCATCCACATGGGCTGCCAGGCTTCCGCTGATCCTTTGAGT 1465
417 lMetTyrThrIleHisMetGlyCysHisGlyPheArgAspProPheGluC 434
1466 GCAACATGTCGCGTACACAGCAGGACCGGTACGAGTTCTCGTCGCAC 1515
434 ysAsnMetCysGlyTyrHisSerGlnAspArgTyrGluPheSerSerHis 450

```

```

1516 ATAACGCGAGGGAGCAGCCGCTTCCACATGAGC 1548
451 lIleThrArgGlyGluHisArgPheHisMetSer 461

```

seq\_name: /cgn2\_6/ptodata/2/paa/US090\_COMB.pap:US-09-019-348A-25

```

seq_documentation_block:
: Sequence 25, Application US/09019348A
: GENERAL INFORMATION:
: APPLICANT: Georgopoulos, Katia
: APPLICANT: Morgan, Bruce A.
: TITLE OF INVENTION: AIOLOS GENE
: FILE REFERENCE: 10287-031001
: CURRENT APPLICATION NUMBER: US/09/019,348A
: CURRENT FILING DATE: 1998-02-05
: PRIOR APPLICATION NUMBER: US 08/733,622
: PRIOR FILING DATE: 1996-10-17
: PRIOR APPLICATION NUMBER: US 60/017,646
: PRIOR FILING DATE: 1996-05-14
: PRIOR APPLICATION NUMBER: US 60/005,529
: PRIOR FILING DATE: 1995-10-18
: NUMBER OF SEQ ID NOS: 38
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 25
: LENGTH: 461
: TYPE: PRT
: ORGANISM: Homo sapiens
: US-09-019-348A-25

```

```

alignment_scores:
Quality: 2467.00      Length: 461
Ratio: 5.351         Gaps: 0
Percent Similarity: 100.000      Percent Identity: 100.000

```

alignment\_block:

US-08-711-417C-165 x US-09-019-348A-25

Align seg 1/1 to: US-09-019-348A-25 from: 1 to: 461

```

166 AATGTTAAAGTAGAGCTCAGAGTCATGAAGAGATGGCGTGTGTA 215
|||||
1 AsnValLysValGluThrGlnSerAspGluGluAsnGlyArgAlaCysGl 17
216 AATGAATGGGGAAGATGTGGGAGGATTTAGCATGCTTGCCTCGG 265
|||||
17 unetAsnGlyGluGluCysAlaGluAspLeuArgMetLeuAspAlaSerG 34
266 GAGAGAAATGAATGGCTCCACAGGGACCAAGGAGCTCGGCTTGTG 315
|||||
34 lyGluLysMetAsnGlySerHisArgAspGlnGlySerSerAlaLeuSer 50
316 GGAGTTGGAGGATTCGACTTCCTTAACGGAAAACTAAAGTGTGATCTG 365
|||||
51 GlyValGlyGlyIleArgLeuProAsnGlyLysLeuLysCysAspIleCy 67
366 TGGGATCATTTGCTATCGGCGCCCAATGTGCTCATGTTCAAAAAGAGCC 415
|||||
67 sGlyIleIleCystIleGlyProAsnValLeuMetValHisLysArgSerH 84
416 ACACGTGGAGAACGCCCTCCAGTGCATCAGTCGCGGGGCTCATTCACC 465
|||||
84 isThrGlyGluArgProPheGlnCysAsnGlnCysGlyAlaSerPheThr 100
466 CAGAAGGGCAACCTGCTCCGGCACATCAAGCTGCTTCCGGGGAGAGCC 515
|||||
101 GlnLysGlyAsnLeuLeuArgHisIleLysLeuHisSerGlyGluLys 117
516 CTTCAATATGCACCTCTGCAACTACGCTGCGCGCGGAGGAGCGCCTCA 565
|||||
117 oPheLysCysHisLeuCysAsnTyrAlaCysArgArgArgAspAlaLeuT 134

```

us-08-711-417c-165.rapm

Wed Aug 28 10:05:37 2002

```

566 CTGGCCACTGAGGACGACCTCCGTTGGTAAACCTCACAATGTGGATAT 615
134 hrGlyHisLeuArgThrHisSerValGlyLysProHisLysCysGlyTyr 150
616 TGTGGCCGAAGCTATAACAGACGACGCTCTTTAGAGGAACAATAAGAGCG 665
151 CysGlyArgSerTyrLysGlnArgThrSerLeuGluHisLysGluAr 167
666 CTGCCACAACACTACTTGGAAAGCATGGCTTCCGGGCACACTGTACCCAG 715
167 gCysHisAsnTyrLeuGluSerMetGlyLeuProGlyThrLeuTyrProV 184
716 TCATTAAAGAACTAAGACAGCTGAAATGGCAGAAAGACCTGTGCAAG 765
184 alileLysGluGluThrLysHisSerGluMetAlaGluAspLysCysLys 200
766 ATAGGATCAGAGAGATCTCTGCTGGACAGACTAGCAAGTATGTGCG 815
201 IleGlySerGluArgSerLeuValLeuAspArgLeuAlaSerAsnValAl 217
816 CAACCTAAGAGCTCTATGCTCAGAAATTTCTTGGGACAAAGGCCCTGT 865
217 aLysArgLysSerSerMetProGlnLysPheLeuGlyAspLysGlyLeuS 234
866 CCGACAGCCCTTACGACAGTGCACCTACGAGAGGAGAGAACCAATGATG 915
234 erAspThrProTyrAspSerAlaThrTyrGluLysGluAsnGluMetMet 250
916 AAGTCCACGTGTGAGGACCAAGCCATCAACAGCCCATCACTACTCGG 965
251 LysSerHisValMetAspGlnAlaIleAsnAsnAlaIleAsnTyrLeuGl 267
966 GCGCAGTCCCTGCGCCCGCTGTGACAGCGCCCGCGGCTTCGAGG 1015
267 yAlaGluSerLeuArgProLeuValGlnThrProGlyGlySerGluV 284
1016 TGGTCCCGGTCTACGACCCGATGTACAGCTGCACAGGCGCTCGGAGGCG 1065
284 alValProValIleSerProMetTyrGlnLeuHisArgArgSerGluGly 300
1066 ACCCGCGCTCCCAACCTCGCCCGCAGACAGCGCGCTGGAGTACCTGCT 1115
301 ThrProArgSerAsnHisSerAlaGlnAspSerAlaValGluTyrLeuLe 317
1116 GCTGCTCTCCAGGCCAAGTGTGTCCTCGGACGCGAGCGCTCCCGGA 1165
317 uLeuLeuSerLysAlaLysLeuValProSerGluArgGluAlaSerProS 334
1166 GCAACAGCTGCCAAGACTCCACGACACCGAGACCAACAGAGGAGCAG 1215
334 erAsnSerCysGlnAspSerThrAspThrGluSerAsnAsnGluGluGln 350
1216 CGCAGCGGTCTTATCTACCTGACCAACACATCGCCGCGCAGCGCAACG 1265
351 ArgSerGlyLeuIleTyrLeuThrAsnHisIleAlaArgArgAlaGlnAr 367
1266 CGTGTGCTCAAGGAGGACCGCGCTCTAGACCTGTGCGCGCGCGCT 1315
367 gValSerLeuLysGluGluHisArgAlaTyrAspLeuLeuArgAlaAlaS 384
1316 CCGAGAACTCGCAGGACGCGCTCCCGGTGGTCAGCACCCAGCGGGAGCAG 1365
384 erGluAsnSerGlnAspAlaLeuArgValValSerThrSerGlyGluGln 400
1366 ATCAAGGTGTACAAAGTGGCAACACTGCCGGGTGCTCTTCCTGGATCAGT 1415
401 MetLysValTyrLysCysGluHisCysArgValLeuPheLeuAspHisVa 417
1416 CATGTACACCATCCACATGGGTGCGCACCGGCTTCGATGATCTTTGAGT 1465
417 lMetTyrThrIleHisMetGlyCysHisGlyPheArgAspProPheGluC 434
1466 GCAACATGTGCGGCTACACAGCCAGCGGTACGAGTCTCTCTCGCAC 1515

```

seq\_name: /cgn2\_6/ptodata/2/paa/US097\_COMB.pep:US-09-755-830-38

seq\_documentation\_block:  
; Sequence 38, Application US/09755830  
; GENERAL INFORMATION:  
; APPLICANT: Georgopoulos, Katia  
; TITLE OF INVENTION: IKAROS REGULATORY ELEMENTS AND USES  
; TITLE OF INVENTION: THEREOF  
; FILE REFERENCE: 10287-067001  
; CURRENT APPLICATION NUMBER: US/09/755,830  
; CURRENT FILING DATE: 2001-01-05  
; PRIOR APPLICATION NUMBER: US 08/283,300  
; PRIOR FILING DATE: 1994-07-29  
; PRIOR APPLICATION NUMBER: US 08/238,212  
; PRIOR FILING DATE: 1994-05-02  
; PRIOR APPLICATION NUMBER: US 08/121,438  
; PRIOR FILING DATE: 1993-09-14  
; PRIOR APPLICATION NUMBER: US 07/946,233  
; PRIOR FILING DATE: 1992-09-14  
; NUMBER OF SEQ ID NOS: 43  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 38  
; LENGTH: 461  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; US-09-755-830-38

alignment\_scores:  
Quality: 2467.00 Length: 461  
Ratio: 5.351 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-08-711-417C-165 x US-09-755-830-38 ..  
Align seg 1/1 to: US-09-755-830-38 from: 1 to: 461

166 AATGTTAAAGTAGAGACTCAGAGTGAAGAGAAATGGCGTGCCTGTGA 215  
|||||  
1 AsnValLysValGluThrGlnSerAspGluGluAsnGlyArgAlaCysG 17  
|||||  
216 AATGAATGGGAAGAATGTCGGAGGATTTACGAATGCTTGATGCTCGG 265  
|||||  
17 uMetAsnGlyGluGluCysAlaGluAspLeuArgMetLeuAspAlaSerG 34  
|||||  
266 GAGAGAAATGAATGGTCCCGACAGGACCAAGCAGCTCGGCTTTGTCG 315  
|||||  
34 lyGluLysMetAsnGlySerHisArgAspGlnGlySerSerAlaLeuSer 50  
|||||  
316 GGAGTTGAGGCAATTCGACTTCCTTAACGGAATAAAGTGTGATATCTG 365  
|||||  
51 GlyValGlyGlyIleArgLeuProAsnGlyLysLeuLysCysAspIleCy 67  
|||||  
366 TGGGATCATTTGCTATCGGGCCCAATGTCTCATGTTTCACAAAAGAAGCC 415  
|||||  
67 sglyIleIleCysIleGlyProAsnValLeuMetValHisLysArgSerH. 84  
|||||  
416 ACACCTGGAGAACGGCCCTTCCAGTGCATCATAGTCGGGGCGCTCATACC 465  
|||||  
84 isThrGlyGluArgProPheGlnCysAsnGlnCysGlyAlaSerPheThr 100  
|||||  
466 CAGAAGGGCAACCTCTCTCCGGCACATCAAGCTGCTATTCGGGGAGAAGCC 515  
|||||  
101 GlnLysGlyAsnLeuArgHisIleLysLeuHisSerGlyGluLysPr 117

```

516 CTTCAAAATGCCACTCTGCAACTACGCTGCCGCGGAGGAGCGCCCTCA 565
117 oPheLysCysHisLeuCysAsnTyrAlaCysArgArgAspAlaLeuT 134
566 CTGCCCACTTGAGCGCAGCTCGTGTGGTAAACCTCACAATGTGGATAT 615
134 hrGlyHisLeuArgThrHisSerValGlyLysProHisLysCysGlyTyr 150
616 TGTGGCGGAAGCTATAAACAGCGAAGCTCTTTAGAGGAACATAAAGAGCG 665
151 CysGlyArgSerTyrLysGlnArgThrSerLeuGluGluHisLysGluAr 167
666 CTGCCCAACTACTTGGAAAGCATGGCTCTCCGGGCACACTGTACCCAG 715
167 gCysHisAsnTyrLeuGluSerMetGlyLeuProGlyThrLeuTyrProV 184
716 TCATTAAGAAGAACTAAGCACAGTGAATGGCAGAAGACCTGTGCAAG 765
184 aLileLysGluGluThrLysHisSerGluMetAlaGluAspLeuCysLys 200
766 ATAGGATCAGAGAGATCTCTGCTGGACAGACTAGCAAGTAAATGTGCG 815
201 ileGlySerGluArgSerLeuValLeuAspArgLeuAlaSerAsnValAl 217
816 CAAAGCTTAAGAGCTCTATGCTCTCAAAATTTCTTGGGCAAGGCGCTGT 865
217 aLysArgLysSerSerMetProGlnLysPheLeuGlyAspLysGlyLeuS 234
866 CCGACACGCCCTACGACAGTGCACAGTACGAGAGGAGAGCAAAATGATG 915
234 erAspThrProTyrAspSerAlaThrTyrGluLysGluAsnGluMetMet 250
916 AAGTCCACGTGTGAGGACCAAGCCATCAACAGCCCATCAACTACTCTGG 965
251 LysSerHisValMetAspGlnAlaIleAsnAsnAlaIleAsnTyrLeuG 267
966 GGCCGAGTCCCTCGGCCGCTGTGTGAGAGCGCCCGGCGGTTCGAGG 1015
267 yAlaGluSerLeuArgProLeuValGlnThrProGlyGlySerGluV 284
1016 TGTGTCGGGTATCAGCCCGATGTACAGCTGCACAGCGCGAGCGCTCCG 1065
284 alValProValIleSerProMetTyrGlnLeuHisArgSerGlyGly 300
1066 ACCCCGCGTCCCAACACACCTCGGCCAGAGACGCGCGCTGGAGTACCT 1115
301 ThrProArgSerAsnHisSerAlaGlnAspSerAlaValGluTyrLeuLe 317
1116 GTGTCTCTCAAGGCCAAGTTGGTCCCTCGAGCGCGAGCGCTCCCGCA 1165
317 uLeuLeuSerLysAlaLysLeuValProSerGluArgGluAlaSerProS 334
1166 GCAACAGCTGCCAAGACTCCACGACCGACCGAGAGCAACAGAGAGCAG 1215
334 erAsnSerCysGlnAspSerThrAspThrGluSerAsnAsnGluGln 350
1216 CGCAGCGTCTTATCTACTCTGACCAACACATCCCGCGAGCGCGCAAG 1265
351 ArgSerGlyLeuIleTyrLeuThrAsnHisIleAlaArgArgAlaGlnAr 367
1266 CGTGTCTCTAAGAGAGAGACCGCGCTACGACCTGCTGCGCGCGCGCT 1315
367 gValSerLeuLysGluGluHisArgAlaTyrAspLeuLeuArgAlaAlaS 384
1316 CCGAGAACTCCAGAGCGCGCTCGCGTGTGTCAGCACCGAGCGGGAGCAG 1365
384 erGluAsnSerGlnAspAlaLeuArgValValSerThrSerGlyGluGln 400
1366 ATGAAGTGTACAAGTCCGAACACTGCCGGGTGCTCTTCTGTGATCAGCT 1415
401 MetLysValTyrLysCysGluHisCysArgValLeuPheLeuAspHisVa 417
1416 CATGTACACCATCCACATGGGCTGCCACGGCTTCCGTGATCCTTTTGAGT 1465

```

```

|||||
417 lmetTyrThrIleHisMetGlyCysHisGlyPheArgAspProPheGluC 434
1466 GCAACATGTGCGGCTACACAGCCAGGACCGGTACGAGTTCTCGTCGCAC 1515
|||||
434 yAsnMetCysGlyTyrHisSerGlnAspArgTyrGluPheSerSerHis 450
1516 ATAACCGAGGAGGAGCACCGCTTCCACATGAGC 1548
|||||
451 IleThrArgGlyGluHisArgPheHisMetSer 461

```

seq\_name: /cgn2\_6/ptodata/2/paa/PCTUS\_COMB.pcp:pct-US99-02559-27

seq\_documentation\_block:  
; Sequence 27, Application PC/TUS9902559  
; GENERAL INFORMATION:  
; APPLICANT: Shiseido Co., Ltd.  
; TITLE OF INVENTION: AIOLOS GENE  
; FILE REFERENCE: 10287/031WO1  
; CURRENT APPLICATION NUMBER: PCT/US99/02559  
; EARLIER FILING DATE: 1999-02-05  
; EARLIER FILING DATE: 1998-02-05  
; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 27  
; LENGTH: 518  
; TYPE: PRT  
; ORGANISM: Mus musculus  
PCT-US99-02559-27

alignment\_scores:  
Quality: 2437.00 Length: 521  
Ratio: 4.913 Gaps: 6  
Percent Similarity: 95.202 Percent Identity: 89.635

alignment\_block:

us-08-711-417c-165 x pct-US99-02559-27 ..

Align seg 1/1 to: PCT-US99-02559-27 from: 1 to: 518

```

1 ATGGATGCTGAGGCGTCAAGACATGCTTTCTCATCAGGGAAGAAAG 50
|||||
1 MetAspValAspGluGlyGlnAspMetSerGlnValSerGlyLysGluSe 17
51 CCCCCCTGTAAAGCATCTCCAGATGAGGGCGATGAGCCCATGCCGATCC 100
|||||
17 rProProValSerAspThrProAspGluGlyAspGluProMetProValP 34
101 CCGAGGACCTCTCCACACCTCGGAGACAGCAAGAGCTCCAAGAGTAC 150
|||||
34 roGluAspLeuSerThrThrSerGlyAlaGlnGlnAsnSerLysSerAsp 50
151 AGAGTCGTGGCCAGTAAAGTAAAGTAGACTCAGAGTGTGAAGAGAA 200
|||||
51 ArgGlyMetAlaSerAsnValLysValGluThrGlnSerAspGluGluas 67
201 TGGCGGTGCTCTGAAATGAATGGGGAAGAAATGTGCGGAGGATTTACGAA 250
|||||
67 nGlyArgAlaCysGluMetAsnGlyGluGluCysAlaGluAspLeuArg 84
251 TGCTTGATGCTCGGAGAGAAATGAATGGCTCCACAGGGACCAAGGC 300
|||||
84 etLeuAspAlaSerGlyGluLysMetAsnGlySerHisArgAspGlnGly 100
301 AGCTCGGCTTTGTGCGGAGTTGAGGACATTTCGACTTCCTTAACGGAAACT 350
|||||
101 SerSerAlaLeuSerGlyValGlyGlyIleArgLeuProAsnGlyLysLe 117
351 AAAGTGTGATATCTGTGGGATCATTTGTCATCGGCGCCCATGTGCTCATGG 400
|||||
117 uLysCysAspIleCysGlyIleValCysIleGlyProAsnValLeuMetV 134

```

Wed Aug 28 10:05:37 2002

```

1295 ACAGACCTGCTGGCGCCGCTCCGAGAACTCCGAGACGCGCTCCCGCTG 1344
||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
431 YrGluValLeuIleuArgAlaAlaSerGluAsnSerGlnAspAlaPheArgVal 447
1345 GTCAGCACACAGCGGGAGCAGATGAAGGTGTACAAAGTCCGAAACACTGCCG 1394
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
448 ValSerThrSerGlyGluGlnLeuLysValTyrLysCysGluHisCysAr 464
1395 GGTGCTCTTCCTGGATCAGCTCATGTACACCATCCACATG.....G 1435
464 GValLeuPheLeuAspHisValMetTyrThrIleHisMetGlyCysHisG 481
1436 GTGTCACACGCTTCCGTGATCTTTTGTAGTGCACATGTGCGGTACCCAC 1485
481 LysCysHisGlyPheArgAspProPheGluCysAsnMetCysGlyTyrHis 497
1486 AGCAGGACCGGTACAGTTCCTGCGCACATAAGCGAGGAGGACACCG 1535
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
498 SerGlnAspArgTyrGluPheSerHisIleThrArgGlyGluHisAr 514
1536 CTTCCACATGAGC 1548
514 gTyrHisLeuSer 518

seq_name: /cgn2.6/ptodata/2/paa/US090_COMB.pep:US-09-019-348-27

seq_documentation_block:
; Sequence 27, Application US/09019348
; GENERAL INFORMATION:
; APPLICANT: Georgopoulos, Katia
; TITLE OF INVENTION: AIOLOS GENE
; FILE REFERENCE: 10287/031001
; CURRENT APPLICATION NUMBER: US/09/019,348
; EARLIER FILING DATE: 1998-02-05
; EARLIER APPLICATION NUMBER: US 08/733,622
; EARLIER FILING DATE: 1996-10-17
; EARLIER APPLICATION NUMBER: US 60/017,646
; EARLIER FILING DATE: 1996-05-14
; EARLIER APPLICATION NUMBER: US 60/005,527
; EARLIER FILING DATE: 1995-10-18
; SOFTWARE: FastSeq for Windows Version 3.0
; NUMBER OF SEQ ID NOS: 29
; SEQ ID NO 27
; LENGTH: 518
; TYPE: PRT
; ORGANISM: Mus musculus
; US-09-019-348-27

alignment_scores:
Quality: 2437.00 Length: 521
Ratio: 4.913 Gaps: 6
Percent Similarity: 95.202 Percent Identity: 89.635

alignment_block:
US-08-711-417C-165 x US-09-019-348-27 ..
Align seg 1/1 to: US-09-019-348-27 from: 1 to: 518

1 ATGGATGCTGACGAGGTCAAGACATGTCTTTCTCATCAGGAAGAAAG 50
||||:||||:||||:||||:||||:||||:||||:||||:||||:
1 MetaspValaspGluGlyGlnAspMetSerGlnValSerGlyLysGluSe 17
51 CCCCCCTGTAAAGCGATCTCCAGATGAGGCGGATGAGCCCATGCCGATCC 100
||||:||||:||||:||||:||||:||||:||||:||||:||||:
17 rProProValSerAspThrProaspGluGlyaspGluProMetProValp 34
101 CCGAGGACCTCTCCACCACCTCCGGGAGGACAGCAAGCTCCAGAGTGAC 150
34 roGluAspLeuSerThrThrSerGlyAlaGlnGlnAsnSerLysSerAsp 50
151 AGAGTCGTGGCCAGTAAATGTTAAAGTAGACTCAGAGTATGATGAAGAA 200

```



## alignment\_block:

US-08-711-417C-165 x US-09-019-348A-27 ..

Align seg 1/1 to: US-09-019-348A-27 from: 1 to: 518

```
1 ATGATGCTGACGAGGTCAAGACATGTTCTTCTCATCATCAGGGAAGAAAG 50
1 MetAspValAspGluGlyGlnAspMetSerGlnValSerGlyLysGlu 17
51 CCCCCCTGTAAGCGATACCTCCAGATGAGCGGCGATGAGCCGATCC 100
17 rProValSerAspThrProAspGluGlyAspGluProMetProValP 34
101 CCGAGGACCTCTCCACACCTCCGGAGGACACCAAGCTCCAAGAGTGAC 150
34 roGluAspLeuSerThrThrSerGlyAlaGlnGlnAsnSerLysSerAsp 50
151 AGAGTCGTGGCCAGTAATGTTAAAGTAGAGACTCAGAGTGATGAGAGAA 200
51 ArgGlyMetAlaSerAsnValLysValGluThrGlnSerAspGluGluAs 67
201 TGGCGCTGCTGTGAATGAATGGGGAAGAATGTCGGAGGATTTACGAA 250
67 nGlyArgAlaCysGluMetAsnGlyGluCysAlaGluAspLeuArgm 84
251 TGCTTGATGCTCGGAGAGAAATGAATGGCTCCACAGGACCAAGGC 300
84 etLeuAspAlaSerGlyGluLysMetAsnGlySerHisArgAspGlnGly 100
301 AGCTCGGCTTTGTCGGAGGTGGAGCATTCGACTTCCTAACGGAAACT 350
101 SerSerAlaLeuSerGlyValGlyGlyIleArgLeuProAsnGlyLysLe 117
351 AAAGTGTGATATCTGGGATCATTTGGCATCGGCCCAATGTGCTCATGG 400
117 uLysCysAspIleCysGlyIleValCysIleGlyProAsnValLeuMetV 134
401 TTCACAAAAGAACCCACATGGAGAACGCCCTTCAGTGCATCAATCAGTGC 450
134 alHisLysArgSerHisThrGlyGluArgProPheGlnCysAsnGlnSer 150
451 GGGGCTCATTTACCCAGAGGCGCACTGCTCCGCGCACATCAAGCTCA 500
151 GlyAlaSerPheThrGlnLysGlyAsnLeuLeuArgHisIleLysLeuHi 167
501 TTCGGGGAGAACCCCTTCAATGCCACCTCTGCAACTACGCTCGCGCC 550
167 sSerGlyGluLysProPheLysCysHisLeuLysCysAsnTyrAlaCysArg 184
551 GGAGGACGCCCTCACTGGCCACCTGAGAGACGACCTCCGTTGGTAAACT 600
184 rArgAspAlaLeuThrGlyHisLeuArgThrHisSerValGlyLysPro 200
601 CACAAATGTGGATATGTGCGCGAAGCTATAACAGCGAAGCTCTTAGA 650
201 HisLysCysGlyTyrCysGlyArgSerTyrLysGlnArgSerSerLeuG 217
651 GGAACATAAGAGCGCTGCGACAACTACTTGGAAAGCATGGCCCTTCGG 700
217 uGluHisLysGluArgCysHisAsnTyrLeuGluSerMetGlyLeuProG 234
701 GCACACTGTACCCAGTCAATAAGAAAGAACTAAGCACAGTGAATGGCA 750
234 ly...ValCysProValIleLysGluGluThrAsnHisAsnGluMetAla 249
751 GAAGACCTGTGAAGTAGATCAGACAGATCTCTCGTCTGCAGACACT 800
250 GluAspLeuCysLysIleGlyAlaGluArgSerLeuValLeuAspArgLe 266
801 AGCAAGTAATGTGCCAAACGTAAGAGCTCTATGCTCAGAAATTTCTTG 850
266 uAlaSerAsnValAlaLysArgLysSerSerMetProGlnLysPheLeuG 283
```

```
851 GGGACAAAGGGCTGTCCGACACAGCCCTACGACAGTGCACGAGTACGAGAG 900
283 lyAspLysCysLeuSerAspMetProTyrAspSerAlaAsnTyrGluLys 299
901 GAGAACGAATGATGAAGTCCACGCTGATGAGCAACGACCATCAACACGC 950
300 Glu...AspMetMetThrSerHisValMetAspGlnAlaIleAsnAsnAl 315
951 CATCAACTACCTGGGGCGGAGTCCCTGCGCCGCTGCTGTCAGACGCC 1000
315 alIeAsnTyrLeuGlyAlaGluSerLeuArgProLeuValGlnThrProp 332
1001 CGGCGGTTCGAGGTGCTCCCGGTATCATCAGCCGATGTACCACTGCAC 1050
332 roGlySerSerGluValValProValIleSerSerMetTyrGlnLeuHis 348
1051 AGG...CGCTCGGAGGCGACCCCGCTCCCAAGGCCAAGTTGGTGCCTCG 1097
349 LysProProSerAspGlyProProArgSerAsnHisSerAlaGlnAsp.. 364
1098 CGCGGTGGAGTACCTGCTGCTCTCTCAAGGCCAAGTTGGTGCCTCG 1147
365 .AlaValAspAsnLeuLeuLeuSerLysAlaLysSerValSerSerG 381
1148 AGCCGAGGCGTCCCGAGCAACAGCTGCCAAGACTCCACGAGACACGAG 1197
381 luArgGluAlaSerProSerAsnSerCysGlnAspSerThrAspThrGlu 397
1198 AGCAACACGAGGAGCGGCGGCTTCTATCTACCTGACCAACACCAT 1247
398 SerAsnAlaGluGluArgSerGlyLeuIleTyrLeuThrAsnHisIl 414
1248 CGCCCGAGCGCGCAAGCG...GTGCTCGCTCAAGGAGGAGGACCGCCCT 1294
414 eAsnProHisAlaArgAsnGlyLeuAlaLeuLysGluGluArgAlaT 431
1295 AGCACTGCTGCGCGCGCTCCGAGAACTCGCAGAGCGGCTCCGCGGTG 1344
431 yrGluValLeuArgAlaAlaSerGluAsnSerGlnAspAlaPheArgVal 447
1345 GTCAGCACCGGAGGAGAGATGAAGTGTACAAAGTGCAGAACACTGCCG 1394
448 ValSerThrSerGlyGluGlnLeuLysValTyrLysCysGluHisCysAr 464
1395 GGTGCTCTTCTGATCAGCTCATGTACACATCCACATG.....G 1435
464 gValLeuPheLeuAspHisValMetTyrThrIleHisMetGlyCysHisG 481
1436 GCTCCGACGCTTCCGTGATCTTTTTCAGTGCACATGTGGCGTACCAC 1485
481 lyCysHisGlyPheArgAspProPheGluCysAsnMetCysGlyTyrHis 497
1486 AGCCAGGACCGGTACGAGTTCGTCTGCACATACGCGAGGAGGAGCAGC 1535
498 SerGlnAspArgTyrGluPheSerSerHisIleThrArgGlyGluHisAr 514
1536 CTTCCACATGAGC 1548
514 gTyrHisLeuSer 518
```

seq\_name: /cqn2\_6/ptodata/2/paa/US097\_COMB.pep:US-09-755-830-40

seq\_documentation\_block:

; Sequence 40, Application US/09755830

; GENERAL INFORMATION:

; APPLICANT: Georgopoulos, Katia

; TITLE OF INVENTION: IKAROS REGULATORY ELEMENTS AND USES

; FILE REFERENCE: 10287-067001

; CURRENT APPLICATION NUMBER: US/09/755,830

; CURRENT FILING DATE: 2001-01-05

; PRIOR APPLICATION NUMBER: US 08/283,300

; PRIOR FILING DATE: 1994-07-29

; PRIOR APPLICATION NUMBER: US 08/238,212  
 ; PRIOR FILING DATE: 1994-05-02  
 ; PRIOR APPLICATION NUMBER: US 08/121,438  
 ; PRIOR FILING DATE: 1993-09-14  
 ; PRIOR APPLICATION NUMBER: US 07/946,233  
 ; PRIOR FILING DATE: 1992-09-14  
 ; NUMBER OF SEQ ID NOS: 43  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 40  
 ; LENGTH: 518  
 ; TYPE: PRT  
 ; ORGANISM: Mus musculus  
 ; US-09-755-830-40

alignment\_scores:  
 Quality: 2437.00 Length: 521  
 Ratio: 4.913 Gaps: 6  
 Percent Similarity: 95.202 Percent Identity: 89.635

alignment\_block:

US-08-711-417C-165 x US-09-755-830-40 ..

Align seg 1/1 to: US-09-755-830-40 from: 1 to: 518

```

1 ATGGATGCTGACGAGGCTCAAGACATGTTCTTCTCATCAGGGAAGAAAG 50
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1 MetAspValaspGluGlnaspMetSerGlnValSerGlyLysGluSe 17
51 CCCCCCTGAAGGATPACTCAGATGAGGGCGATGAGCCCATGCCGATCC 100
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
17 rProProValSerAspThrProAspGluGlyAspGluProMetProValP 34
101 CCAGGACCTCTCCACCACTCGGGAGGACAGCAAGCTCCAAAGATGAC 150
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
34 roGluaspLeuSerThrThrSerGlyAlaGlnGlnasnSerLysSerasp 50
151 AGAGTGTGCTGCCAGTAAATGTAAGTAGAGACTCAGAGTGATCAAGAGAA 200
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
51 ArgGlyMetAlaSerAsnValLysValGluThrGlnSerAspGluGluAs 67
201 TGGGCGTGTGTAATGAATGGGGAAGAAATGTGGGAGGATTTACGAA 250
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
67 nGlyArgAlaCysGluMetasnGlyGluGluCysAlaGluaspLeuArgM 84
251 TGCTTGATGCTCGGAGAGAAATGAATGGCTCCACAGGGACCAAGGC 300
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
84 etLeuaspAlaSerGlyGluLysMetasnGlySerHisargaspGlnGly 100
301 AGCTCGGCTTTGTGGGAGTTGGAGCATTCGACTTCCTTAACGGAAACT 350
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
101 SerSerAlaLeuSerGlyValGlyIleArgLeuProAsnGlyLysLe 117
351 AAAGTGTGATCTGTGGGATCATTTGTCATCGGGCCCAATGTCTCATGG 400
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
117 uLysCysaspIlecysGlyIleValGlyIleGlyProasnValLeuMetV 134
401 TTCACAAAAGAACCACTGGAGAGCGGCCCTTCCAGTGCAATCATGTC 450
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
134 alHisLysArgSerHisThrGlyGluArgProPheGlnCysAsnGlnSer 150
451 GGGGCTCATTCACCCAGAGGGCAACCTGCTCCGGCACATCAAGCTGCA 500
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
151 GlyAlaSerPheThrGlnLysGlyAsnLeuLeuArgHisIleLysLeuH 167
501 TTCGGGGAGAGCCCTTCAATGCCACCTCTCACTACGCTCGCCGCC 550
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
167 sSerGlyGluLysProPheLysCysHisLeuCysasnThrAlaCysArgA 184
551 GGAGGGAGCCCTCACATGGCCACCTGAGGACGCACTCCGTTGGTAAACCT 600
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
184 rArgaspAlaLeuThrGlyHisLeuArgThrHisSerValGlyLysPro 200

```

```

601 CACAAATGTGGATATTGTGGCCGAGAGCTATAAACACGACGACGCTCTTTAGA 650
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
201 HisLysCysGlyTyrCysGlyArgSerTyrLysGlnArgSerSerLeuG 217
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
651 GGAACATAAGAGCGCTGCCACAACTACTTGGAAAGCATGGGCTTCGG 700
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
217 uGluHisLysGluArgCysHisasnTyrLeuGluSerMetGlyLeuProG 234
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
701 GCACACTGTACCCAGTCATTAAAGAAAGAACTAAGCACAGTGAATGGCA 750
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
234 ly...ValCysProValIleLysGluGluThrAsnHisAsnGluMetAla 249
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
751 GAAGACCTGTGCAAGATAGGATCAGACAGATCTCTGCTGCTGCACAGACT 800
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
250 GluaspLeuCysLysIleGlyAlaGluArgSerLeuValLeuaspArgLe 266
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
801 AGCAAGTAATGTGCGCAACAGTAAGAGCTCTATGCTCTCAGAAAATTTCTTG 850
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
266 uAlaSerAsnValAlaLysArgLysSerSerMetProGlnLysPheLeuG 283
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
851 GGGACAAAGGCCCTGTCCGACACGCCCTTACGACAGTGCACGTACGAGAG 900
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
283 lyAspLysCysLeuSerAspMetProTyrAspSerAlaAsnTyrGluLys 299
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
901 GAGAACAAATGATGAAGTCCCAGCTGATGACCAAGCCCATCAACACGC 950
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
300 Glu...AspMetMetThrSerHisValMetaspGlnAlaIleasnAla 315
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
951 CATCAACTACCTGGGGCCGAGTCCCTCGCGCCGCTGGTGACAGCCTCC 1000
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
315 alLeasnTyrLeuGlyAlaGluSerLeuArgProLeuValGlnThrProp 332
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1001 CGGCGGTTCGAGGTGTCGGGTATCATCAGCCCGCATGTACCAGCTGCAC 1050
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
332 roGlySerSerGluValProValIleSerSerMetTyrGlnLeuHis 348
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1051 AGG...CGCTCGAGGGCACCCCGCTGCCAACCACTCGGCCACGAGCAG 1097
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
349 LysProSerAspGlyProProArgSerAsnHisSerAlaGlnasp... 364
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1098 CGCGGTGGAGTACCTGCTGCTCTCCAAGGCCAAGTTGGTCCCTCGG 1147
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
365 .AlaValaspAsnLeuLeuSerLysAlaLysSerValSerSerG 381
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1148 AGCGGGAGGCTCCCGAGCAACAGCTCCCAAGACTCCACGACACCCGAG 1197
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
381 luArgGluAlaSerProSerAsnSerCysGlnaspSerThrAspThrGlu 397
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1198 AGCAACAACAGGAGCAGCGAGCGGTCTTATCTACCTGACCAACACCAT 1247
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
398 SerAsnAlaGluGluArgSerGlyLeuIleTyrLeuThrAsnHisIl 414
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1248 CGCCCGAGCGCGGCAACGC...GTGCTCCTCAAGGAGGACGACCGCCT 1294
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
414 eAsnProHisAlaArgasnGlyLeuAlaLeuLysGluGluArgAla 431
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1295 AGGACCTGCTCGCGCCCGCTCCGAGAACTCGCAGGCGCTCGCGCTG 1344
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
431 yrGluValLeuArgAlaAlaSerGluasnSerGlnaspAlaPheargVal 447
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1345 GTCAGCACACGCGGGAGCAGATGAAGTGTACAAAGTCCGAACACTGCCG 1394
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
448 ValSerThrSerGlyGluGlnLeuLysValTyrLysCysGluHisCysAr 464
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1395 GTGTCTCTCTCGGATCAGCTCATCTACACCATCCACATG.....G 1435
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
464 gValLeuPheLeuAspHisValMetTyrThrIleHisMetGlyCysHisG 481
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1436 GCTGCCAGCGCTTCGTGATCTCTTTGAGTGCAACATGTGGGCTACAC 1485
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
481 lycyHisHisGlyPheargaspProPheGluCysasnMetCysGlyTyrHis 497
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1486 AGCCAGGACCGGTACGAGTTCTCTGTCGACATACGCGAGGGGACCG 1535

```

498 SerGlnAspArgTyrGluPheSerHisIleThrArgGlyGluHisAr 1514  
 1536 CTTCCACATGAGC 1548  
 1548 TTTTCTTTT  
 514 gTyrHisLeuSer 518

seq\_name: /cgn2\_6/ptodata/2/paa/PCTUS\_COMB.pcp: PCT-US99-04224-7

seq\_documentation\_block:

; Sequence 7, Application PC/TUS9904224A  
 ; GENERAL INFORMATION:  
 ; APPLICANT: The General Hospital Corporation  
 ; TITLE OF INVENTION: THE HELIOS GENE  
 ; FILE REFERENCE: 10287/043W01  
 ; CURRENT APPLICATION NUMBER: PCT/US99/04224A  
 ; CURRENT FILING DATE: 1999-02-26  
 ; EARLIER APPLICATION NUMBER: US 60/076,325  
 ; EARLIER FILING DATE: 1998-02-27  
 ; NUMBER OF SEQ ID NOS: 17  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 7  
 ; LENGTH: 517  
 ; TYPE: PRT  
 ; ORGANISM: Mus musculus  
 PCT-US99-04224-7

alignment\_scores:

Quality: 2426.50 Length: 521  
 Ratio: 4.892 Gaps: 7  
 Percent Similarity: 95.202 Percent Identity: 89.443

alignment\_block:

US-08-711-417C-165 x PCT-US99-04224-7 ..

Align seg 1/1 to: PCT-US99-04224-7 from: 1 to: 517

1 ATGATGCTGACGAGGTGACACATGCTCTTCTCATCAGGGAGAAAG 50  
 1 MetAspValAspGluGlyGlnAspMetSerGlnValSerGlyLysGluSe 17  
 51 CCCCCCTGTAAAGCATCTCCACATGAGGGCGATGAGCCCATCCCATCC 100  
 17 rProProValSerAspThrProAspGluGlyAspGluProMetProValP 34  
 101 CCGAGGACCTTCACACACTCCGGGAGGACGACCAAGCTCCACAGAGTGAC 150  
 34 roGluAspLeuSerThrThrSerGlyAlaGlnGlnAsnSerLysSerAsp 50  
 151 AGAGTCTGCGCCAGTAACTTAACTAGAGACTCAGAGTATGAGAGAA 200  
 51 ArgGlyMetGlySerAsnValLysValGluThrGlnSerAspGluGluAs 67  
 201 TGGCGCTGCTGTGAATGAATGGGAAGAATGTGGGAGGATTTACGAA 250  
 67 nGlyArgAlaCysGluMetAsnGlyGluGluCysAlaGluAspLeuArgW 84  
 251 TGCTGATGCTCGGAGAGAAATGAATGGCTCCACAGGACCAAGGC 300  
 84 etLeuAspAlaSerGlyGluLysMetAsnGlySerHisArgAspGlnGly 100  
 301 AGCTCGGCTTTGTCGGAGTGGAGGCATTCGATCTCCATCAAGGAAACT 350  
 101 SerSerAlaLeuSerGlyValGlyGlyIleArgLeuProAsnGlyLysLe 117  
 351 AAAGTGTGATATGTGGGATCATTTGGCATCGGGCCCAATGTGCTCATGG 400  
 117 uLysCysAspIleCysGlyIleValCysIleGlyProAsnValLeuMetV 134  
 401 TTCACAAAAGACCCACTGGAGACGCGCCCTTCCAGTGCATCAGTGC 450  
 134 alHisLysArgSerHisThr...GluArgProPheGlnCysAsnGlnCys 149

451 GGGGCTCTTCCACCCAGAGGCAACCTGCTCCGGCACATCAAGCTGCA 500  
 150 GlyAlaSerPheThrGlnLysGlyAsnLeuLeuArgHisIleLysLeuH 166  
 501 TTCCGGGAGAGCCCTTCAAAATGCCACCTCTGCAACTACGCTGCGCGC 550  
 166 sSerGlyGluLysProPheLysCysHisLeuCysAsnTyrAlaCysArg 183  
 551 GGAGGACGCCCTCACTGCGCCACCTGAGGACGCACTCCGTTGGTAAACCT 600  
 183 rgArgAspAlaLeuThrGlyHisLeuArgThrHisSerValGlyLysPro 199  
 601 CACAAATGTGGATATGTGCGCGCAAGCTATAACAGCAGCACTCTTTAGA 650  
 200 HisLysCysGlyTyrCysGlyArgSerTyrLysGlnArgSerSerLeuG 216  
 651 GGAACATAAAGAGCGCTGCCACAACTACTTGGAAAGCATGGCGCTTCGG 700  
 216 uGluHisLysGluArgCysHisAsnTyrLeuGluSerMetGlyLeuProG 233  
 701 GCACACTGTACCCAGTCAATTAAGAAGAACTAACACAGTGAATGGA 750  
 233 ly...ValCysProValIleLysGluGluThrAsnHisAsnGluMetAla 248  
 751 GAAGACCTGTGCAAGATAGGATCAGAGAGATCTCTCGTCTGGACAGACT 800  
 249 GluAspLeuCysLysIleGlyAlaGluArgSerLeuValLeuAspArg 265  
 801 AGCAAGTATGTCGCCAACGTAAGAGCTCTATGCTCAGAAATTTCTTG 850  
 265 uAlaSerAsnValAlaLysArgLysSerSerMetProGlnLysPheLeuG 282  
 851 GGGACAAGGGCTGTCGACAGCGCTACGACAGCTGCGCCACCTACGAGAG 900  
 282 lyAspLysCysLeuSerAspMetProTyrAspSerAlaAsnTyrGlnLys 298  
 901 GAGAACGAAATGATGAGTCCCACGTGATGGACCAAGCCATCAACAAGC 950  
 299 Glu...AspMetMetThrSerHisValMetAspGlnAlaIleAsnAsnAl 314  
 951 CATCAACTACCTGGGGCGGAGTCCCTCGCGCCGCTGGTGAGAGCGCCC 1000  
 314 alLeuAsnTyrLeuGlyAlaGluSerLeuArgProLeuValGlnThrProp 331  
 1001 CCGGCGGTTCGAGGTGTCGCGGTATCATCGCCGATGTACCCAGCTGCAC 1050  
 331 roGlySerSerGluValValProValIleSerSerMetTyrGlnLeuHis 347  
 1051 AGG...CGCTCGAGGGCACCCCGCTCCCAACCACTCGGCCAGGACAG 1097  
 348 LysProProSerAspGlyProProArgSerAsnHisSerAlaGlnAsp... 363  
 1098 CCGCGTGGAGTACCTGCTGCTGCTCTCAAGGCCAAGTTGGTCCCTCGG 1147  
 364 AlaValAspAsnLeuLeuLeuSerLysAlaLysSerValSerSerG 380  
 1148 AGCGGAGGCGTCCCGAGCAACAGCTGCCAAGACTCCAGGACACCGAG 1197  
 380 luArgGluAlaSerProSerAsnSerCysGlnAspSerThrAspThrGlu 396  
 1198 AGCAACACGAGGAGCAGCGCGGTCTTATCTACCTGACCAACACAT 1247  
 397 SerAsnAlaGluGluGlnArgSerGlyLeuIleTyrLeuThrAsnHisI 413  
 1248 CGCCCGAGCGCGCAACGC...GTGTCGCTCAAGGAGGACACCGCGCT 1294  
 413 eAsnProHisAlaArgAsnGlyLeuAlaLeuLysGluGlnArgAla 430  
 1295 ACAGCTGCTGCGCGCGCTCCGAGAACTCGCAGGACGCGCTCCGCGTG 1344  
 430 yrGluValLeuArgAlaAlaSerGluAsnSerGlnAspAlaPheArgVal 446

```
1345 GTCAGCACCGGGAGGAGCAGATGAAGTGTACAAAGTGCAGGAACACTGCCG 1394
|||||
447 ValSerThrSerGlyGluGlnLeuLysValTyrLysCysGluHisCysAr 463
|||||
1395 GGTGCTCTTCCTGATCAGTCATGTCATACACATCCACATG.....G 1435
|||||
463 gValLeuPheLeuAspHisValMetTyrThrIleHisMetGlyCysHisG 480
|||||
1436 GCTGCCACCGGCTTCGGTGCATCTCTTCAGTGCACATGTCGGGTACCAC 1485
|||||
480 lYcysHisGlyPheArgAspProPheGluCysAsnMetCysGlyTyrHis 496
|||||
1486 AGCCAGGACCGGTACGAGTCTCTCGTCGCACATAACGGGAGGGAGCACCG 1535
|||||
497 SerGlnAspArgTyrGluPheSerSerHisIleThrArgGlyGluHisAr 513
|||||
1536 CTTCCACATGAGC 1548
|||||
513 gTyrHisLeuSer 517

seq_name: /cgn2_6/ptodata/2/paa/US092_COMB.pap:US-09-259-389-7

seq_documentation_block:
; Sequence 7, Application US/09259389
; GENERAL INFORMATION:
; APPLICANT: Georgopoulos, Katia
; APPLICANT: Morgan, Bruce A.
; APPLICANT: Kelly, Clair
; TITLE OF INVENTION: THE HELIOS GENE
; FILE REFERENCE: 10287/043001
; CURRENT APPLICATION NUMBER: US/09/259,389
; CURRENT FILING DATE: 1999-02-26
; EARLIER APPLICATION NUMBER: US 60/076,325
; EARLIER FILING DATE: 1998-02-27
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 517
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-259-389-7
```

```
alignment_scores:
  Quality: 2426.50      Length: 521
  Ratio: 4.892          Gaps: 7
Percent Similarity: 95.202 Percent Identity: 89.443
```

alignment\_block:

US-08-711-417c-165 x US-09-259-389-7 ..

Align seg 1/1 to: US-09-259-389-7 from: 1 to: 517

```
1 ATCGATGCTGACGAGGTCACAGCATGTCTTCTCATCAGGGAAGAAAG 50
|||||
1 MetAspValAspGluGlnAspMetSerGlnValSerGlyLysGluSe 17
|||||
51 CCCCCCTGTAAAGCGATCTCCAGATGAGGGCGATGAGCCCATGCCGATCC 100
|||||
17 rProValSerAspThrProAspGluGlyAspGluProMetProValP 34
|||||
101 CCAGGACCTCTCCACCTCGGGAGGACAGCAAGCTCCAGAGTGAC 150
|||||
34 roGluAspLeuSerThrThrSerGlyAlaGlnAsnSerLysSerAsp 50
|||||
151 AGAGTCGTGGCCAGTAAATGTTAAAGTAGAGACTCAGAGTGATCAAGAGAA 200
|||||
51 ArgGlyMetGlySerAsnValLysValGluThrGlnSerAspGluGluAs 67
|||||
201 TGGCGTGCCTGTGAATGAATGGGAAGAAATGTCGGGAGGATTTACGAA 250
|||||
67 nGlyArgAlaCysGluMetAsnGlyGluGluCysAlaGluAspLeuArgM 84
|||||
```

```
251 TCGCTGATCCCTCGGAGAGAAAATGAATGGCTCCACAGGACCAAGGC 300
|||||
84 etLeuAspAlaSerGlyGluLysMetAsnGlySerHisArgAspGlnGly 100
|||||
301 AGCTCGGCTTTCGGGAGTTCGAGGCATTTCGACTTCCTTAACGGAACACT 350
|||||
101 SerSerAlaLeuSerGlyValGlyGlyIleArgLeuProAsnGlyLysLe 117
|||||
351 AAAGTGTGATATCTGTGGGATCATTTTCATCGGGCCCAATGTGCTCATGG 400
|||||
117 uLysCysAspIleCysGlyIleValCysIleGlyProAsnValLeuMetV 134
|||||
401 TTCACAAAAGAACCCACACTGGAGAGCGCCCTTCACGTGCATCATGTCG 450
|||||
134 alHisLysArgSerHisThr...GluArgProPheGlnCysAsnGlnCys 149
|||||
451 GGGGCTCATTTACCCAGAGGCAACCTGTCTCCGGCAGCATCAAGCTGCA 500
|||||
150 GlyAlaSerPheThrGlnLysGlyAsnLeuLeuArgHisIleLysLeuHi 166
|||||
501 TTCGGGGAGAGCCCTTCAATGCCACCTCTGCAACTACGCTGCCGCC 550
|||||
166 sSerGlyGluLysProPheLysCysHisLeuCysAsnTyrAlaCysArgA 183
|||||
551 GGAGGAGCGCCCTCACTGGCCACCTGAGGACGCACTCCGTTGGTAAACCT 600
|||||
183 rgArgAspAlaLeuThrGlyHisLeuArgThrHisSerValGlyLysPro 199
|||||
601 CACAAATCTGGATATTGTGGCCGAAGCTATAAACAGCAGCACTTTTAGA 650
|||||
200 HisLysCysGlyTyrCysGlyArgSerTyrLysGlnArgSerSerLeuGl 216
|||||
651 GGAACATAAAGAGCGCTGCCAACACTACTTGGAAAGCATGGGCTTCCGG 700
|||||
216 uGluHisLysGluArgCysHisAsnTyrLeuGluSerMetGlyLeuPro 233
|||||
701 GCACACTGTACCCAGTCATTAAAGAGAACTAAGCACAGCTGAAATGGCA 750
|||||
233 ly...ValCysProValIleLysGluThrAsnHisAsnGluMetAla 248
|||||
751 GAAGACCTGTGCAAGATAGGATCAGAGATCTCTCGTGTGGACAGACT 800
|||||
249 GluAspLeuCysLysIleGlyAlaGluArgSerLeuValLeuAspArgLe 265
|||||
801 AGCAAGTAATGTCGCCAAGCTAAGAGCTCTATGCCTCAGAAATTTCTTG 850
|||||
265 uAlaSerAsnValAlaLysArgLysSerSerMetProGlnLysPheLeuG 282
|||||
851 GGGACAAAGGCGCTCTCCGACACGCGCTTACGACAGTGCACGTACGAGAAG 900
|||||
282 lYAspLysCysLeuSerAspMetProTyrAspSerAlaAsnTyrGluLys 298
|||||
901 GAGAACGAATGATGAAGTCCACGTCATGGACCAAGCCATCAACAACGC 950
|||||
299 Glu...AspMetMetThrSerHisValMetAspGlnAlaIleAsnAl 314
|||||
951 CATCAACTACCTCGGGCGGAGTCCCTCGCGCCGCTGTGTGAGACGCGCCC 1000
|||||
314 alLeasnTyrLeuGlyAlaGluSerLeuArgProLeuValGlnThrProp 331
|||||
1001 CGGCGGTTCGAGGTGTCCTCGGTCTATCAGCCCGATGTACCACTGCAC 1050
|||||
331 roGlySerSerGluValValProValIleSerSerMetTyrGlnLeuHis 347
|||||
1051 AGG...CGCTCGGAGGCGCCCGCTCCCAACCACTCGGCCAGGACAG 1097
|||||
348 LysProProSerAspGlyProProArgSerAsnHisSerAlaGlnAsp.. 363
|||||
1098 CGCCGTGGAGTACCTGTCTGCTGCTCTCCAAAGCCCAAGTTGTCGCCCTCG 1147
|||||
364 .AlaValAspAsnLeuLeuLeuSerLysAlaLysSerValSerSerG 380
|||||
1148 AGCGGAGGCGGTCCCGGAGAGCAACAGCTCTCCAAAGACTCCACGACACCGAG 1197
|||||
```

```

|||||
380 luArgGluAlaSerProSerAsnSerCysGlnAspSerThrAspThrGlu 396
1198 AGCAACACGAGGAGGAGCGCGGTCTTATCTACCTGACCAACACAT 1247
|||||
397 SerAsnAlaGluGluArgSerGlyLeuIleThrLeuThrAsnHis11 413
1248 CGCCCGACGCGCGCAACGC...GTGTCGCTCAAGGAGGAGCAGCGGCGCT 1294
|||||
413 eAsnProHisAlaArgAsnGlyLeuAlaLeuLysGluGluArgAlaT 430
1295 ACACCTGCTGCGCGCGCTCGGAGAACTCGGAGGAGCGGCTCGCGTG 1344
|||||
430 yrGluValLeuArgAlaAlaSerGluAsnSerGlnAspAlaPheArgVal 446
1345 GTCAGCACGAGCGGAGGAGGAGTGTACAACTGCGAACTGCGCG 1394
|||||
447 ValSerThrSerGlyGluGlnLeuLysValTyrLysCysGluHisCysAr 463
1395 GGTGCTCTCTCGGATCAGTCATGTACACCATCCACATG.....G 1435
|||||
463 gValLeuPheLeuAspHisValMetTyrThrIleHisMetGlyCysHisG 480
1436 GCTGCCACGCTCCGCTGATCTCTTTGAGTGCAACATGTGCGGCTACAC 1485
|||||
480 lYcHisGlyPheArgAspProPheGluCysAsnMetCysGlyTyrHis 496
1486 AGCAGGACGCTGACGATGTCGTCGACATACGCGAGGAGGAGCAGCG 1535
497 SerGlnAspArgTyrGluPheSerSerHisIleThrArgGlyGluHisAr 513
1536 CTTCCACATGAGC 1548
513 gtyrHisLeuSer 517

```

```
seq_name: /cgn2_6/ptodata/2/paa/US081_COMB.pep.US-08-121-438-5
```

```

seq_documentation_block:
; Sequence 5, Application US/08121438
; GENERAL INFORMATION:
; APPLICANT: Georgopoulos, Katia A.
; TITLE OF INVENTION: IKAROS: A T CELL PATHWAY REGULATORY GENE
; NUMBER OF SEQUENCES: 152
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 STATE STREET, Suite 510
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/121,438
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 946,233
; FILING DATE: 09-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Paul L.
; REGISTRATION NUMBER: 35,695
; REFERENCE/DOCKET NUMBER: MPG-006
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 568 amino acids
; TYPE: amino acid

```

```

; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
; US-08-121-438-5

alignment_scores:
  Quality: 2422.00      Length: 571
  Ratio: 4.863         Gaps: 7
  Percent Similarity: 87.215  Percent Identity: 82.137

alignment_block:
US-08-711-417C-165 x US-08-121-438-5

Align seg 1/1 to: US-08-121-438-5 from: 1 to: 568

1 ATGGATGCTGACGAGGTCAAGACATGCTTTCTCATCAGGGAAGAAAG 50
|||||
1 MetAspValAspGluGlyGlnAspMetSerGlnValSerGlyLysGluSe 17
|||||
51 CCCCCCTGTAAAGCATACTCCAGATGAGGGGATGAGCCCATCCGATCC 100
|||||
17 rProProValSerAspThrProAspGluGlyAspGluProMetProValP 34
|||||
101 CCGAGGACCTCTCCACCACCTCGGGAGGACGAAAGCTCCAAGAGTGAC 150
|||||
34 roGluAspLeuSerThrThrSerGlyAlaGlnGlnAsnSerLysSerAsp 50
|||||
151 AGAGTCGTGGCCAGTAATGTTAAAGTAGAGACTCAGAGTGATGAAGAGAA 200
|||||
51 ArgGlyMetAlaSerAsnValLysValGluThrGlnSerAspGluGluAs 67
201 TGGCGTGCCTGTGAATGAATGCGGAAGAATGTGCGGAGGATTTACGAA 250
|||||
67 nGlyArgAlaCysGluMetAsnGlyGluGluCysAlaGluAspLeuArgM 84
251 TGCTTGATGCTCGGGAGAGAAAATGAATGGCTCCACAGGAGGACCAAGGC 300
|||||
84 etLeuAspAlaSerGlyGluLysMetAsnGlySerHisArgAspGlnGly 100
300 .....
101 SerSerAlaLeuSerGlyValGlyGlyIleArgLeuProAsnGlyLysLe 117
300 .....
117 uLysCysAspIleCysGlyIleValCysIleGlyProAsnValLeuMetV 134
300 .....
134 aHisLysArgSerHisThrGlyGluArgPropheGlnCysAsnGlnCys 150
301 AGCTCGGCTTTGTCGGAGTGGAGGCATTCGACTTCCTTAACGGAACACT 350
|||||
151 SerSerAlaLeuSerGlyValGlyGlyIleArgLeuProAsnGlyLysLe 167
351 AAAGTGTGATATCTGTGGGATCATTTGTCATCGGGCCCAATGTCTCATCG 400
|||||
167 uLysCysAspIleCysGlyIleValCysIleGlyProAsnValLeuMetV 184
401 TTCACAAAAGAGCCACACTGGAGAACGCGCTTCCAGTGCATCAGTGCC 450
|||||
184 aHisLysArgSerHisThrGlyGluArgPropheGlnCysAsnGlnCys 200
451 GGGGCTCATTTCCACCAGAGGGCAACCTGTCTCCGGCAGCATCAAGCTGCA 500
|||||
201 GlyAlaSerPheThrGlnLysGlyAsnLeuLeuArgHisIleLysLeuHi 217
501 TTCCGGGAGAGCGCTTCAAAATGCCACCTCTGCAACTACGCTGCCGCC 550
|||||
217 sSerGlyGluLysProPheLysCysHisLeuCysAsnTyrAlaCysArgA 234

```

```

551 GGAGGAGCGCCCTCAGTGGCCACCTGAGGACGACCTCCGTTGGTAAACCT 600
234 r9argaspAlaLeuThrGlyHisLeuArgThrHisSerValGlyLysPro 250
601 CACAAATGTGGATATTGTGGCCGAAGCTATAACACAGCGAACCTCTTTAGA 650
251 HisLysCysGlyTyrCysGlyArgSerTyrLysGlnArgSerSerLeuG1 267
651 GGRACATAAGAGCGCTGCCACACACTPACTTGGAAAGCATGGCCCTCCGG 700
267 uGluHisLysGluArgCysHisAsnTyrLeuGluSerMetGlyLeuProG 284
701 GCACACTGTACCCAGTCATTAAAGAGAACTAAACACAGCTGAATGGCA 750
284 ly...MetTyrProValIleLysGluGluThrAsnHisGluMetAla 299
751 GRAGACTGTGCAAGATAGATCAGAGATCTCTCGTGTGGACAGACT 800
300 GluAspLysCysLysIleGlyAlaGluArgSerLeuValLeuAspArgLe 316
801 AGCAAGTAATGTCCGCAACGCTAAGAGCTCTATGCCCTCAGAAATTTCTG 850
316 uAlaSerAsnValAlaLysArgLysSerSerMetProGlnLysPheLeuG 333
851 GCGACAAGGCGCTGCCGACAGCGCCTACGACAGTGCACGCTACGAGAAG 900
333 lyAspLysCysLeuSerAspMetProTyrAspSerAlaAsnTyrGluLys 949
901 GAGAACGAAATGATGAAGTCCACGCTGATGACCAAGCCATCAACAGCG 950
350 Glu...AspMetMetThrSerHisValMetAspGlnAlaIleAsnAl 365
951 CATCAACTACTCGGGGCGAGTCCCTCGCGCGCTGCTGCACAGCGCCC 1000
365 alIeAsnTyrLeuGlyAlaGluSerLeuArgProLeuValGlnThrProp 382
1001 CGGCGGTTCGAGGTGTCGCCGTCTCCAGGCTACAGCCCGATGTACCACTGCAC 1050
382 roGlySerSerGluValProValIleSerSerMetTyrGlnLeuHis 398
1051 AGG...CGCTCGGAGGCGACCCCGCGCTCCCAACCACTCGCGCCGACGAC 1097
399 LysProProSerAspGlyProProArgSerAsnHisSerAlaGlnAsp.. 414
1098 CGCGGTGAGTACCTGCTGCTGCTCCAGGCGCAAGGTTGGTGCCTCGG 1147
415 .AlaValAspAsnLeuLeuLeuSerLysAlaLysSerValSerSerG 431
1148 AGCGGAGGCGTCCCGAGCAACAGCTGCCAAGACTCCACAGGACACCGAG 1197
431 luArgGluAlaSerProSerAsnSerCysGlnAspSerThrAspThrGlu 447
1198 AGCAACAACGAGGAGGACGAGCGCGTCTTATCTACCTGACCAACACAT 1247
448 SerAsnAlaGluGlnArgSerGlyLeuIleTyrLeuThrAsnHisI 464
1248 CGCCCGACGCGCGCAACG...GTCTGCTCAAGGAGGACCGCGCCT 1294
464 eAsnProHisAlaArgAsnGlyLeuAlaLeuLysGluGluArgAla 481
1295 AGCACTGCTGCGCGCGCTCCGAGAACTCGCAGGACGCGCTCCGCGTG 1344
481 yGluValLeuArgAlaAlaSerGluAsnSerGlnAspAlaPheArgVal 497
1345 GTCAGCACGAGCGGAGCAGATGAAGTGTACAAAGTGCAGCACTGCCG 1394
498 ValSerThrSerGlyGluGlnLeuLysValTyrLysCysGluHisCysAr 514
1395 GTGCTCTTCTGATCAGCTGATGTACACCATCCACATG.....G 1435
514 gValLeuPheLeuAspHisValMetTyrThrIleHisMetGlyCysHisG 531
1436 GCTGCGCGGCTCCGATGATCCCTTTGAGTGCAACATGTGCGGCTACCA 1485

```

```

|||||
531 lyCysHisGlyPheArgAspPropGluCysAsnMetCysGlyTyrHis 547
1486 AGCCAGGACCGGTACGAGTTCGTCGCGCACATAACCGGAGGAGCACCG 1535
|||||
548 SerGluAspArgTyrGluPheSerSerHisIleThrArgGlyGluHisAr 564
1536 CTCCACATGAGC 1548
|||||
564 gTyrHisLeuSer 568

seq_name: /cgn2_6/ptodata/2/paa/PCTUS_COMB.pcp.pct-US99-02559-22
seq_documentation_block:
; Sequence 22, Application PC/TUS9902559
; GENERAL INFORMATION:
; APPLICANT: Shiseido Co., Ltd.
; TITLE OF INVENTION: AIOLOS GENE
; FILE REFERENCE: 10287/031W01
; CURRENT APPLICATION NUMBER: PCT/US99/02559
; CURRENT FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: US 09/019,348
; EARLIER FILING DATE: 1998-02-05
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 22
; LENGTH: 470
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(470)
; OTHER INFORMATION: xaa = Any Amino Acid
PCT-US99-02559-22

```

## alignment\_scores:

```

Quality: 2207.50      Length: 468
Ratio: 5.098          Gaps: 3
Percent Similarity: 92.521 Percent Identity: 90.385

```

## alignment\_block:

```

US-08-711-417C-165 x PCT-US99-02559-22 ..

```

```

Align seg 1/1 to: PCT-US99-02559-22 from: 1 to: 470

```

```

160 GCAGTAATGTTAAAGTAGAGACTCAGAGTGATGAGAGAAATGGCGTGC 209
|||||
3 AlaserAsnValLysValGluThrGlnSerAspGluGluAsnGlyArgAl 19
210 CTGTGAATGAATGGGGAAGAATGTGCGGAGGATTTACGAATGCTTGATG 259
|||||
19 acysGluMetAsnGlyGluGluCysAlaGluAspLeuArgMetLeuAspA 36
260 CCTCGGAGAGAAAATGAATGCTCCACAGGACGACCAAGCAGCTCGCT 309
|||||
36 laseGlyGluLysMetAsnGlySerHisArgAspGlnGlySerSerAla 52
310 TTGTCGGAGTTGGAGGCAATTCGACTTCTTAACGGAACCTAAAGTGTA 359
|||||
53 LeuSerGlyValGlyGlyLeuArgLeuProAsnGlyLysLeuLysCysAs 69
360 TATCTGTGGGATTTGTCATCGGCGCCCAATGTGCTCATGTTCACAAAA 409
|||||
69 pIleCysGlyIle**CysIleGlyProAsnValLeuMetValHisLysA 86
410 GAAGCACTGGAGAACGGCCCTTCCAGTGCAATCAGTCCGCGGCGCTCA 459
|||||
86 rgSerHisThrGlyGluArgProPheGlnCysAsnGlnCysGlyAlaSer 102
460 TTCACCCAGAGGCAACCTGCTCCCGGACATCAACGCTCATTTCCGGGGA 509
|||||
103 PheThrGlnLysGlyAsnLeuLeuArgHisIleLysLeuHisSerGlyCl 119

```

```

alignment_scores:
  Quality: 2207.50
  Ratio: 5.098
  Percent Similarity: 92.521
  Length: 468
  Gaps: 3
  Percent Identity: 90.385

alignment_block:
  US-08-711-417C-165 x US-08-238-212A-153 ..
  Align seg 1/1 to: US-08-238-212A-153 from: 1 to: 470

```

510 GAAGCCCTTCAAAATGCCACCTCTGCAACTACGCCCTGCCGCCGAGGAGC 559  
|||||  
119 uIysProPheLysCysHisLeuCysAsnTyraLaCysArgArGAspa 136  
|||||  
560 CCTCACTGGCCACCTGAGGAGCCTCGTTGGTAAACCTCACAAATGT 609  
|||||  
136 laLeuThrGlyHisLeuArgThrHisSerValGlyLysProHisLysCys 152  
|||||  
610 GGATATTGGCCGAAGCTATAAACAGCAGCAACCTCTTTAGAGAAACATA 659  
|||||  
153 GlyTyrcysGlyArgSerTyrlsGlnArg\*\*\*SerLeuGluGluHisLy 169  
|||||  
660 AGAGCGCTGCCAACACTACTTGGAAAGCATGGCCCTCCGGGCACACTGT 709  
|||||  
169 sGluArGcYsHisAsnTyrlEuGluSerMetGlyLeuProGlyLys\*\*\*\* 186  
|||||  
710 ACCAGTCAATTAAGAGAAGAACTAAGCACAGTAGTGAATGGCAGAGACTG 759  
|||||  
186 \*\*ProValIleLysGluGluThr\*\*\*His\*\*\*GluMetalGluAspLeu 202  
|||||  
760 TGCAAGATAGGATCAGAGAGATCTCTCGTCTGGCAGACTAGCAAGTAA 809  
|||||  
203 CysLysIleGly\*\*\*GluArGserLeuValLeuAspArgLeuAlaSerAs 219  
|||||  
810 TGTGCGCAACGTAAGAGCTCTATGCCCTCAAAATTTCTTGGGACACAGG 859  
|||||  
219 nValAlaLysArgLysSerSerMetProGlnLysPheLeuGlyAspLys\* 236  
|||||  
860 GCCTGTCCGACACGCCCTACGACAGTGCACCTACGAGAGGAGACGAA 909  
|||||  
236 \*\*LeuSerAsp\*\*\*ProTyraSpSerAla\*\*\*TyrlGluLysGlu\*\*\*\*\* 252  
|||||  
910 ATGATGAAGTCCGAGTGTATGACCAAGCAACATCAACAACGCCATCAACTA 959  
|||||  
253 MetMet\*\*\*SerHisValMetAsp\*\*\*AlaIleAsnAlaIleAsnTy 269  
|||||  
960 COTGGGGCGGAGTCCCTCGCGCCGCTGGTGGAGAGCGCCCGCGCGTT 1009  
|||||  
269 rLeuGlyAlaGluSerLeuArgProLeuValGlnThrProProGlyLys\*\*S 286  
|||||  
1010 CGAGGTGGTCCCGGTCTATCAGCCCGATACAGCTGACAGTGCAC...AGGCGC 1056  
|||||  
286 erGluValValProValIleSerProMetTyrlGlnLeuHis\*\*\*\*\* 302  
|||||  
1057 TCGGAGGGCACCCCGCGCTCCAAACACTCGGCCAGGACAGCGCCGTGA 1106  
|||||  
303 Ser\*\*\*Gly\*\*\*ProArgSerAsnHisSerAlaGlnAsp\*\*\*AlaVal\*\* 319  
|||||  
1107 GTPACTGTGCTGCTCTCCAGGCCCAAGTTGGTGCCTCGGAGCGCGAG 1156  
|||||  
319 \*\*\*\*LeuLeuLeuSerLysAlaLys\*\*\*Val\*\*\*SerGluArgGluA 336  
|||||  
1157 CGTCCCGAGCAACAGCTGCCAAGATCCACGAGACACCGAGAGCAACAC 1206  
|||||  
336 laSerProSerAsnSerCysGlnAspSerThrAspThrGluSerAsn\*\* 352  
|||||  
1207 GAGAGCAGCGCAGCGTCTATCTACTACCAACACACATCGCCCGCAGC 1256  
|||||  
353 GluGluGlnArgSerGlyLeuIleTyrlEuThrAsnHisIle\*\*\*\*\* 369  
|||||  
1257 CGCG...CAACGCGTGTGCTCAAGGAGGAGCACCGCGCTACGACCTGC 1303  
|||||  
369 \*Ala\*\*\*\*\*LeuLysGluGlu\*\*\*ArgAlaTyrl\*\*\*\*\*L 386  
|||||  
1304 TGGCGGCGCGCTCCGAGAACTCGAGGACCGCGTCCGCGTGGTCAGACAC 1353  
|||||  
386 euAr-gAlaAlaSerGluAsnSerGlnAspAla\*\*\*ArgValValSerThr 402  
|||||  
1354 AGCGGGGAGCAGATGAAGGTGTACAAGTGCAGACACTCCCGGGTGCTCTT 1403  
|||||  
403 SerGlyGluGlnp\*\*\*rIvsValTyrlLysCysGluHisCysArgValLeuPh 419  
|||||

```

160 GCCAGTAATCTTAAAGTAGAGACTCAGAGTGTAGTGAAGAGAAATGGCGTGC 209
161 |||||
162 3 AlaSerAsnValLysValGluThrGlnSerAspGluGluAsnGlyArgAl 19
163 |||||
164 210 CTGTGAATGAATGGGAAAGAAATGTGGGAGGATTTAGCAATGCTTGATG 259
165 |||||
166 19 aCysGluMetAsnGlyGluGluCysAlaGluAspLeuArgMetLeuAspA 36
167 |||||
168 260 CTTCCGGAGAGAAATGAATGCTCCCAAGGACCAAGGAGGAGTTCGGCT 309
169 |||||
170 36 laSerGlyGluLysMetAsnGlySerHisArgAspGlnGlySerSerAla 52
171 |||||
172 310 TTGTCCGGAGTTGGAGGATTCGACTTCCTAACGGAACAACTAAAGTGTGA 359
173 |||||
174 53 LeuSerGlyValGlyGlyLeuArgLeuProAsnGlyLysLeuLysCysAs 69
175 |||||
176 360 TATCTGTGGGATCATTTGATCGGGCCCAATGTCTCATGTTTCCAAAA 409
177 |||||
178 69 pileCysGlyIle***CysIleGlyProAsnValLeuMetValHisLysA 86
179 |||||
180 410 GAAGCCACACTGGAGAACGGCCCTTCCAGTGCATCATGTCGGGGCTCA 459
181 |||||
182 86 rgSerHisThrGlyGluArgProPheGlnCysAsnGlnCysGlyAlaSer 102
183 |||||
184 460 TTCACCCAGAGAGGCAACCTGCTCCGGACACATCAAGCTGCATTCCGGGA 509
185 |||||
186 103 PheThrGlnLysGlyAsnLeuLeuArgHisIleLysLeuHisSerGlyGI 119
187 |||||
188 510 GAAGCCCTTCAATGCCACCTCTGCACTAGCGCTCGCCCGGAGGAGC 559
189 |||||
190 119 uLysProPheLysCysHisLeuLysCysAsnTyrAlaCysArgArgAspA 136
191 |||||
192 560 CCCTCACTGGCCACTGAGGACGACCTCCGTTGGTAAACCTCAACAATGT 609
193 |||||
194 136 laLeuThrGlyHisLeuArgThrHisSerValGlyLysProHisLysCys 152
195 |||||
196 610 GGATATTGTGGCCGAGCTATAACAGAGCAACGCTTTTAGAGAAACATAA 659
197 |||||
198 153 GlyTyrCysGlyArgSerTyrLysGlnArg***SerLeuGluGluHisLy 169
199 |||||
200 560 AGAGCGTGCACCAACTACTTGTGAAGCATGGCCCTTCGGGGCACACTGT 709
201 |||||
202 169 sGluArgCysHisAsnTyrLeuGluSerMetGlyLeuProGly***** 186
203 |||||
204 710 ACCAGTCAATTAAGAGAACTAAGCAGCTGAAATGGCAGAACCTG 759
205 |||||
206 186 **ProValIleLysGluGluThr***His**GluMetAlaGluAspLeu 202
207 |||||
208 760 TGCAAGATAGGATCAGAGAGATCTCTCGTGGGACAGACTAGCAAGTAA 809
209 |||||
210 203 CysLysIleGly**GluArgSerLeuValLeuAspArgLeuAlaSerAs 219
211 |||||
212 810 TGTCGCCAACTAGAGCTCTATCGCTCAGAAATTTCTGGGGACAAGG 859
213 |||||
214 219 nValAlaLysArgLysSerSerMetProGlnLysPheLeuGlyAspLys* 236
215 |||||
216 860 GCGTGTCCGACAGCCCTACGACAGTGCAGCTAGCAGAGAGGAAAGAA 909
217 |||||
218 236 **LeuSerAsp***ProTyrAspSerAla***TyrGluLysGlu***** 252
219 |||||
220 910 ATGATGATGATCCAGCTGATGGACCAAGCCATCAACACGCTCAACTA 959
221 |||||
222 253 MetMet***SerHisValMetAsp**AlaIleAsnAlaIleAsnTy 269
223 |||||
224 960 CCTGGGGCCGAGTCCCTGCGCCCTGCTGGTGCAGACGCCCGCGGGCGTT 1009
225 |||||
226 269 rLeuGlyAlaGluSerLeuArgProLeuValGlnThrProProGly***S 286
227 |||||
228 1010 CCGAGGTGGTCCCGGTCAATCAGCCCGGATGATACAGCTGCAC...AGCGC 1056
229 |||||
230 286 erGluValValProValIleSerProMetTyrGlnLeuHis***** 302

```

```

1057 TCGAGGGGCAACCCCGCGCTCCCAACCACTCGGCCAGGACAGCGCGTGA 1106
1058 |||||
1059 303 Ser***Gly***ProArgSerAsnHisSerAlaGlnAsp***AlaVal** 319
1060 |||||
1061 1107 GTACCTCTGCTGCTCCCAAGGCCAAGTTGGTGCCTCGGAGCGCGAGG 1156
1062 |||||
1063 319 ***LeuLeuLeuSerLysAlaLys***Val***SerGluArgGluA 336
1064 |||||
1065 1157 CTTCCCGGAGCAACAGCTGCCAAGACTCCAGGACCGAGAGCAACAAC 1206
1066 |||||
1067 336 laSerProSerAsnSerCysGlnAspSerThrAspThrGluSerAsn*** 352
1068 |||||
1069 1207 GAGGAGCAGCGAGCGGCTTATCTACTGACCAACACACATCGCCCGAGC 1256
1070 |||||
1071 353 GluGluGlnArgSerGlyLeuIleTyrLeuThrAsnHisIle***** 369
1072 |||||
1073 1257 CCGC...CAACGCGTGTGCTCAAGGAGGAGCAGCGCTACGACCTGC 1303
1074 |||||
1075 369 *Ala*****LeuLysGluGlu***ArgAlaTyr*****L 386
1076 |||||
1077 1304 TCGCGCGCGCTCCGAGAACTCCGAGACGGCTCCGCGTGGTGCAGCAC 1353
1078 |||||
1079 386 euArgAlaIleSerGluAsnSerGlnAspAla***ArgValValSerThr 402
1080 |||||
1081 1354 AGCGGGGAGCAGATGAAGGTGACAAGTGCAGACACTGCGCGGTGCTCT 1403
1082 |||||
1083 403 SerGlyGluGln**LysValTyrLysCysGluHisCysArgValLeuPh 419
1084 |||||
1085 1404 CTTGATCAGCTCATGTACACCATTCACATG.....GGCTGCCAG 1444
1086 |||||
1087 419 eLeuAspHisValMetTyrThrIleHisMet*****GlyCysHisG 436
1088 |||||
1089 1445 GCTTCCGTGATCTTTTGTAGTGCAACATGTGCGGCTACCAAGCCAGGAC 1494
1090 |||||
1091 436 lyPheArgAspProPheGluCysAsnMetCysGlyTyrHisSerGlnAsp 452
1092 |||||
1093 1495 CGGTACAGTTCGTCGCACATACGCGAGGGGAGCAGCGCTTCCACAT 1544
1094 |||||
1095 453 ArgTyrGluPheSerSerHisIleThrArgGlyGluHisArg***His** 469
1096 |||||
1097 1545 GAGC 1548
1098 |||||
1099 469 *Ser 470

```

```
seq_name: /cgn2_6/ptodata/2/paa/US084_COMB.pep:US-08-465-590B-153
```

```

seq_documentation_block:
; Sequence 153, Application US/08465590B
; Patent No. 5777072
; GENERAL INFORMATION:
; APPLICANT: Georgopoulos, Katia A.
; TITLE OF INVENTION: IKAROS: A T CELL PATHWAY REGULATORY GENE
; NUMBER OF SEQUENCES: 191
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; CITY: BOSTON
; STREET: 28 STATE STREET, Suite 510
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Ascii (text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/465,590B
; FILING DATE: 05-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/238,212
; FILING DATE: 02-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/121,438

```

Wed Aug 28 10:05:37 2002

```

; FILING DATE: 14-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/946,233
; FILING DATE: 14-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Paul L. 35,695
; REGISTRATION NUMBER:
; REFERENCE/DOCKET NUMBER: MPG-006C2DV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 153:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 470 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: C-terminal
;
; US-08-465-590B-153

```

```

alignment_scores:
  Quality: 2207.50      Length: 468
  Ratio: 5.098          Gaps: 3
  Percent Similarity: 92.521  Percent Identity: 90.385

```

```

alignment_block:
US-08-711-417C-165 x US-08-465-590B-153  ..
Align seg 1/1 to: US-08-465-590B-153  from: 1 to: 470

```

```

160 GCCAGTAATGTTAAAGTAGAGACTCAGAGTCATGAAGAGAAATGGCGTGC 209
3 AlaSerAsnValLysValGluThrGlnSerAspGluGluAsnGlyArgAla 19
210 CTGCTGAATGAATGGGAGAGATCGCGAGGATTTACGAATGCTTGATG 259
19 acysGluMetAsnGlyGluGluCysAlaGluAspLeuArgMetLeuAspA 36
260 CCTCGGAGAGAAAATGAATGCTGCCACAGGGACCAAGCAGCAGCTCGGCT 309
36 laSerGlyGluLysMetAsnGlySerHisArgGlnGlySerSerAla 52
310 TTGCTGGGAGTTGGAGGATTCGACTTCTTAACGGAAAACATAAGTGTGA 359
53 LeuSerGlyValGlyGlyLeuArgLeuProAsnGlyLysLeuLysCysAs 69
360 TATCTGTGGGATCATTTGCATCGGCGCCCAATGTGCTCATGTTCAACAAA 409
69 pIleCysGlylle**CysIleGlyProAsnValLeuMetValHisLysA 86
410 GAAGCCACACTGGAGAACGGCCCTTCCAGTGCATCAAGTGGCGGCTCA 459
86 rgSerHisThrGlyGluArgProPheGlnCysAsnGlnCysGlyAlaSer 102
460 TTCACCCAGAGGCAACCTGCTCCGGCACATCAAGCTGCATTCGGGGA 509
103 PheThrGlnLysGlyAsnLeuLeuArgHisIleLysLeuHisSerGlyGl 119
510 GAAGCCCTTCAATGCCACTCTGCAACTACGCCCTGCCCGCGGAGGAGC 559
119 uLysProPheLysCysHisLeuLysAsnTyrAlaCysArgArgArgAspA 136
560 CCCTCACTGGCCACCTGAGGACCACTCGGTTGGTAAACCTCACAAATGT 609
136 laLeuThrGlyHisLeuArgThrHisSerValGlyLysProHisLysCys 152
610 GGATATTGTGGCCGAAGCTATAACAGCGAACCTCTTTAGAGGAACATAA 659
153 GlyTyrCysGlyArgSerTyrLysGlnArg***SerLeuGluGluHisLy 169
660 AGAGCGCTGCCACAACACTACTTGGAAAGCATGGCGCTTCGCGGCACACTGT 709

```

```

|||||
169 sGluArgCysHisAsnTyrLeuGluSerMetGlyLeuProGly***** 186
710 ACCAGTCATTAAAGAAAGAACTAAGCACAGTGAATGGCAGAGACCTG 759
186 **ProValIleLysGluGluThr***His***GluMetAlaGluAspLeu 202
760 TCAGATAGATAGATCAGAGATCTCTGCTGTGGACAGACTAGCAAGTAA 809
203 CysLysIleGly***GluArgSerLeuValLeuAspArgLeuAlaSerAs 219
810 TGTCCGCAAAACGTAAGAGCTCTATGCCCTCAGAAATTTCTTGGGACAAG 859
219 nValAlaLysArgLysSerSerMetProGlnLysPheLeuGlyAspLys* 236
860 GCGTGTCCGACACGCCCTACGACATGCGCAGCTAGCAGAGAGAACGAA 909
236 **LeuSerAsp***ProTyrAspSerAla***TyrGluLysGlu***** 252
910 ATGATGAAGTCCCGAGTGTGATGGACCAAGCCATCAACAGCCCATCACTA 959
253 MetMet***SerHisValMetAsp***AlaIleAsnAsnAlaIleAsnTy 269
960 CTTGGGGGGCGAGTCCCTCGCGCCGCTGTGCAGACGCCCGCGGCGGTT 1009
269 rLeuGlyAlaGluSerLeuArgProLeuValGlnThrProProGly***S 286
1010 CCGAGGTGTCCCGGTCTATCAGCCCGATGTACAGCTGCAC...AGGCGC 1056
286 erGluValValProValIleSerProMetTyrGlnLeuHis***** 302
1057 TCGAGGGCACCCCGCGCTCCCAACACTCGGCCAGGACAGCGCGTGGGA 1106
303 Ser***Gly***ProArgSerAsnHisSerAlaGlnAsp***AlaVal** 319
1107 GTACTCTGCTGCTCTCTCCAAGGCCAAGTTGGTGGCCCTCGGAGCGGAG 1156
319 ***LeuLeuLeuLeuSerLysAlaLys***Val***SerGluArgGluA 336
1157 GTCCTCCGAGCAACAGCTGCAAGACTCCACGACACACCGAGAGCAACAAC 1206
336 laSerProSerAsnSerCysGlnAspSerThrAspThrGluSerAsn*** 352
1207 GAGGAGCAGCGCGGTCTTATCTACTGACCAACCATCGCCGCGCAGC 1256
353 GluGluGlnArgSerGlyLeuIleTyrLeuThrAsnHisIle***** 369
1257 CCGC...CAACGCGTGTGCTCAAGGAGGAGCAGCGCGCTACGACCTGC 1303
369 *Ala*****LeuLysGluGlu***ArgAlaTyr*****L 386
1304 TCGCGCGCGCTCCGAGAACTCGCAGGACGGCTCCGCGTGGTCAGCAC 1353
386 euArgAlaAlaSerGluAsnSerGlnAspAla***ArgValValSerThr 402
1354 AGCGGGAGCAGATGAAGTGTACAAAGTGCAGAACACTGCCGGGTGCTCTT 1403
403 SerGlyGluGln***LysValTyrLysCysGluHisCysArgValLeuPh 419
1404 CCTGGATACGTCATGTACACCATCCACATG.....GGCTGCCACG 1444
419 eLeuAspHisValMetTyrThrIleHisMet*****GlyCysHisG 436
1445 GCTTCGCTGATCCTTTTGTAGTCAACATGTGGCGCTACACAGCAGGAC 1494
436 lypheArgAspProPheGluCysAsnMetCysGlyTyrHisSerGlnAsp 452
1495 CGGTACGAGTCTCTGTCGACATAACCGCAGGGGAGCAGCCGCTCCACAT 1544
453 ArgTyrGluPheSerSerHisIleThrArgGlyGluHisArg***His** 469
1545 GAGC 1548
|||

```

469 \*Ser 470

seq\_name: /cgn2\_6/ptodata/2/paa/US087\_COMB.pep:US-08-733-622A-22

seq\_documentation\_block:

; Sequence 22, Application US/08733622A  
; GENERAL INFORMATION:  
; APPLICANT: Katia Georgopoulos  
; TITLE OF INVENTION: The Aiolos Gene  
; NUMBER OF SEQUENCES: 22  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD, LLP  
; STREET: 28 State Street  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02109-1775

; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/733.622A  
; FILING DATE: 17-OCT-1996

; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: 60/017,646  
; FILING DATE: 14-MAY-1996

; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: 60/005,529  
; FILING DATE: 18-OCT-1995

; ATTORNEY/AGENT INFORMATION:  
; NAME: Myers, Louis  
; REGISTRATION NUMBER: 35,965  
; REFERENCE/DOCKET NUMBER: MGP-042CP  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617)227-7400  
; TELEFAX: (617)742-4214

; INFORMATION FOR SEQ ID NO: 22:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 470 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear

; MOLECULE TYPE: peptide  
; FRAGMENT TYPE: C-terminal  
; US-08-733-622A-22

## alignment\_scores:

Quality: 2207.50 Length: 468  
Ratio: 5.098 Gaps: 3  
Percent Similarity: 92.521 Percent Identity: 90.385

## alignment\_block:

US-08-711-417c-165 x US-08-733-622A-22 ..

Align seg 1/1 to: US-08-733-622A-22 from: 1 to: 470

160 GCCAGTAATGTTAACTAGACTCAGACTGATGAGAGAAATGGCGGTGC 209  
3 AlasSerAsnValLysValGluThrGlnSerAspGluGluAsnGlyArgAl 19  
210 CTCTGAAATGAATGGGAAAGATGTGCGGAGGATTTACGAATGCTTGATG 259  
19 acYsGluMetAsnGlyGluGluCysAlaGluAspLeuArgMetLeuAsp 36  
260 CTTCCGGGACAGAAATGATGCTCCACAGGACCAAGGACGAGCTCGGCT 309  
36 laSerGlyGluLysMetAsnGlySerHisArgAspGlnGlySerSerAla 52  
310 TTGTCGGGAGTTGGAGGCAATTCGACTTCCTAACGGAANAACCTAAAGTGTA 359

53 LeuSerGlyValGlyGlyIleArgLeuProAsnGlyLysLeuLysCysAs 69  
360 TATCTGTGGATCATTTGCATCGGGCCCAATGTGCTCATGTTTCACAAA 409  
69 pileCysGlyIle\*\*\*CysIleGlyProAsnValLeuMetValHisLys 86  
410 GAAGCCACACTGGAGACGGCCCTTCAGTGCATCAATCAGTCCGGGGCTCA 459  
86 rgSerHisThrGlyGluArgProPheGlnCysAsnGlnCysGlyAlaSer 102  
460 TTCACCCAGAAAGGCAACCTGCTCCGGCAGATCAAGCTGCATTCCGGGGA 509  
103 PheThrGlnLysGlyAsnLeuLeuArgHisIleLysLeuHisSerGlyG 119  
510 GAAGCCCTTCAATGCCACCTCTGCACTAGCGCTCCCGCCGGAGGAGC 559  
119 uLysProPheLysCysHisLeuCysAsnTyrAlaCysArgArgAsp 136  
560 CCCTCACTGCCACCTGAGGACGACTCCGTTGGTAAACCTCACAAATGT 609  
136 laLeuThrGlyHisLeuArgThrHisSerValGlyLysProHisLysCys 152  
610 GGATATTGTGGCGAAGCTATAAACAGACGCAAGCTCTTTAGAGAAATAA 659  
153 GlyTyrCysGlyArgSerTyrLysGlnArg\*\*\*SerLeuGluGluHisL 169  
660 AGAGCGCTGCCACACTACTTGGAAAGCATGGGCTTCGGGCGACACTGT 709  
169 sGluArgCysHisAsnTyrLeuGluSerMetGlyLeuProGly\*\*\*\* 186  
710 ACCAGTCATTAAAGAAAGAACTAAGCAGTGAATGGCAGAAAGACCTG 759  
186 \*\*ProValIleLysGluGluThr\*\*\*His\*\*\*GluMetAlaGluAspLeu 202  
760 TCAAGATAGGATCAGAGATCTCTGCTGGACAGACTAGCAAGTAA 809  
203 CysLysIleGly\*\*\*GluArgSerLeuValLeuAspArgLeuAlaSerAs 219  
810 TGTGCCCAAGCTAGAGCTCTATGCTCAGAAATTTCTTGGGCAAGG 859  
219 nValAlaLysArgLysSerSerMetProGlnLysPheLeuGlyAspLys 236  
860 GCCTGTCCGACACGCCCTAGCAGTCCACGTACGAGAGAGAGAGAGAA 909  
236 \*\*LeuSerAsp\*\*\*ProTyrAspSerAla\*\*\*TyrGluLysGlu\*\*\*\* 252  
910 ATGATGAAGTCCACGTGATGGACCAAGCCATCAACAACGCCATCAACTA 959  
253 MetMet\*\*\*SerHisValMetAsp\*\*\*AlaIleAsnAlaIleAsnTy 269  
960 CTTGGGGCCGAGTCCCTGCGCCGCTGCTGTCAGACGCCCGCGGGGTT 1009  
269 rLeuGlyAlaGluSerLeuArgProLeuValGlnThrProProGly\*\*\*S 286  
1010 CCGAGTGTGTCCTCGCTCATCAGCCCGATGTACAGCTGCAC...AGGCGC 1056  
286 erGluValValProValIleSerProMetTyrGlnLeuHis\*\*\*\*\* 302  
1057 TCGGAGGCGCCCGCGCTCCCAACCACTCGGCCAGGACAGCGCGCTGGA 1106  
303 Ser\*\*\*Gly\*\*\*ProArgSerAsnHisSerAlaGlnAsp\*\*\*AlaVal\*\* 319  
1107 GTACCTGCTGCTCTCCAGGCCCAAGTTGGTCCCTCGGAGCGCGAGG 1156  
319 \*\*\*LeuLeuLeuSerLysAlaLys\*\*\*Val\*\*\*SerGluArgGlu 336  
1157 CTTCCCGGACGAACTGCAAGTCCACGAGTCCACGACACCGAGAGCAACAAC 1206  
336 laSerProSerAsnSerCysGlnAspSerThrAspThrGluSerAsn\*\*\* 352  
1207 GAGGAGCAGCGCGGTCTTATCTACTGACCAACCAACATCGCCGACG 1256

Wed Aug 28 10:05:37 2002

```

353 GluGluGlnArgSerGlyLeuIleTyrLeuThrAsnHisIle***** 369
1257 CGCG...CAACGGGTGCGCTCAAGAGGACACCGCGCTACGACGTGC 1303
||| ::::::::::::::::::::|||
369 *Ala*****LeuLySGluLu***ArgAlaTyr*****L 386
13304 TGCSCCGCGCTCCGAGAAGCTCGAGAGCGCGCTCCGCGTGGTCAGCAC 1353
||| ::::::::::::::::::::|||
386 euArgAlaLaSerGluAsnSerGluAspAla***ArgValValSerThr 402
13354 AGCGGGAGCAGATGAAGGTGTACAGTGCAGAACACTGCGCGGTGCTCTT 1403
403 SerGlyGluGln***LysValTyrLysCysGluHisCysArgValLeuPh 419
1404 CCTGGATCAGTGTATGACACCATCCACATG.....GGCTGCCAGC 1444
||| ::::::::::::::::::::|||
419 eLeuAspHisValMetTyrThrIleHisMet*****GlyCysHisG 436
1445 GCTTCGCGTATCCCTTTTGTAGTGCACATGTGCGGCTACACAGCCAGGAC 1494
||| ::::::::::::::::::::|||
436 LysPheArgAspProPheGluCysAsnMetCysGlyTyrHisSerGlnAsp 452
1495 CGGTACGAGTTCCTCGTCGCACATACCGGAGGAGCACCGCTTCACAT 1544
||| ::::::::::::::::::::|||
453 ArgTyrGluPheSerSerHisIleThrArgGlyGluHisArg***His* 469
1545 GAGC 1548
|||
469 *Ser 470

```